



UNIVERSITÉ D'ORLÉANS



ÉCOLE DOCTORALE SCIENCES ET TECHNOLOGIES

Station de Radioastronomie de Nançay

THÈSE présentée par :

Dalal AIT ALLAL

soutenue le : **16 Novembre 2012**

pour obtenir le grade de : **Docteur de l'université d'Orléans**

Discipline : **Automatique et Traitement de Signal**

**Détection d'évènements impulsionnels en
environnement radioélectrique perturbé.
Application à l'observation des pulsars
intermittents avec un système temps réel de
traitement du signal**

THÈSE DIRIGÉE PAR :

Gilles THEUREAU

Astronome, Station de radioastronomie de Nançay

RAPPORTEURS :

Aris KARASTERGIOU

Research Fellow, Université d'Oxford - Angleterre

David SMITH

Directeur de Recherche, C.E.N. de Bordeaux-Gradignan

JURY :

Alain BAUDRY

Professeur, Université Bordeaux 1

Nicole CORNILLEAU-WEHRLIN

Directrice de recherche, Ecole Polytechnique

Rachid ARBA

Professeur, Université d'Orléans

Aris KARASTERGIOU

Research Fellow, Université d'Oxford - Angleterre

David SMITH

Directeur de Recherche, C.E.N. de Bordeaux-Gradignan

Gilles THEUREAU

Astronome, Station de radioastronomie de Nançay

Remerciements

Je remercie d’abord mon directeur de thèse Gilles Theureau et mes encadrants Rodolphe Weber et Ismael Cognard, qui ont toujours su me soutenir et m’encourager durant ces quatre années de thèse. Je tiens aussi à remercier Nicole Cornilleau-Wehrlin, directrice de recherche au CNRS, qui était ma directrice de thèse pendant mes deux premières années de thèse.

Je suis très reconnaissante à messieurs David Smith, directeur de recherche au CNRS, Centre d’Etudes Nucléaires de Bordeaux Gradignan, et Aris Karastergiou Research Fellow à l’université d’Oxford, d’avoir jugé mon travail et d’en avoir été les rapporteurs. J’associe également l’ensemble du jury à mes remerciements.

Je tiens à remercier le personnel de la station de radioastronomie de Nançay pour son accueil chaleureux. Je m’associe également à la Direction de la station de radioastronomie de Nançay pour remercier le Conseil Régional de la Région Centre de son support financier.

Mes remerciements s’adressent aussi aux directeurs des laboratoires qui m’ont accueilli, Monsieur Michel Tagger directeur du LPC2E et Monsieur Christophe Léger responsable de l’équipe “Images et Signaux pour les Systèmes” à l’institut PRISME. Je tiens à remercier également les personnes qui m’ont aidée de près ou de loin tout au long de ce travail. Mes pensées vont à Rémy Leconge, Rym Féliachi, Cédric Dumez-Viou pour leur disponibilité et leur soutien. Je n’oublie pas Meryem Jabloun pour le remontage de moral d’urgence au laboratoire.

Je termine mes remerciements par ceux à qui je dédie cette thèse : mes parents, mes deux sœurs et mon âme sœur. Je remercie aussi ma meilleure amie qui m’a soutenue dans les moments difficiles.

Table des matières

Introduction	9
1 L'observation radio des pulsars	11
1.1 Introduction	11
1.2 L'instrumentation pour l'observation des pulsars	12
1.2.1 Le radiotélescope de Nançay	12
1.2.2 La dispersion par le milieu interstellaire	12
1.2.3 La dédispersion incohérente	14
1.2.4 La dédispersion cohérente	14
1.2.5 L'algorithme de l'empilement 'Folding'	15
1.2.6 L'architecture et les ressources	15
1.3 Modèle du signal d'un pulsar	16
1.4 L'impact des RFIs sur la qualité des observations	19
2 Traitement des interférences pour les observations de pulsars	21
2.1 Introduction	21
2.2 Introduction générale sur les interférences radioélectriques	21
2.3 Détecteurs d'interférences fondés sur le critère de puissance	23
2.3.1 Détecteur de radar	23
2.3.2 Seuillage combinatoire : SumThreshold	24
2.3.3 Filtrage "zéro DM"	24
2.3.4 conclusions	26
2.4 CycloDet : Détecteur de brouilleurs cyclostationnaires	28
2.4.1 Présentation du concept	28
2.4.2 Exemple de la modulation BPSK :	29
2.4.3 Les autres modulations :	29
2.4.4 CycloDet et CycloDet 'aveugle' :	30
2.4.5 Conclusions	31
2.5 Estimateur spectral du Kurtosis	31
2.6 Evaluation des techniques pour le chronométrage des pulsars	33
2.6.1 Robustesse des détecteurs à la présence d'un pulse de pulsar	33
2.6.2 Sensibilité des détecteurs en présence de RFI de type BPSK	38
2.6.3 Sensibilité des détecteurs après dédispersion	40
2.7 L'influence du filtrage sur les performances de détecteurs	42
2.8 Exemple d'implantation matériel : Uniboard	51
2.9 Application sur des données réelles	53
2.10 Conclusion et perspectives	56
3 Détection d'événements impulsionnels intermittents	59
3.1 Introduction	59
3.1.1 Détection d'impulsions individuelles avec PRESTO	60
3.2 Détection de pulses individuelles avec SIPSFAR	60
3.2.1 Principe de SIPSFAR	60
3.2.2 La transformée de Radon	63
3.2.3 Comparaison de la sensibilité	64
3.3 Résultats obtenues sur des données du pulsar du Crabe	67

3.4	Conclusion	69
4	Analyse du survey FOSTER	71
4.1	Introduction	71
4.2	Le programme de recherche de nouveaux pulsars 'FOSTER'	71
4.3	Recherche d'événements impulsionnels par la méthode SIPSFAR dans un GPU	71
4.3.1	Architecture du GPU et la librairie CUDA	72
4.3.2	Application aux données dites 'FOSTER'	73
4.4	Comparaisons des résultats SIPSFAR et PRESTO	74
4.4.1	Une observation très parasitée	75
4.4.2	Le pulsar du Crabe B0531+21	76
4.4.3	Le pulsar PSR J0248+6021	77
4.4.4	Un traitement statistique simple	80
4.5	Conclusion	85
	Conclusion	87
	A La radio cognitive	89
	B Impact du filtre de mise en forme	91
	C La transformée de Radon	93
	D Calcul des sensibilités $S_{SIPSFAR}$ et S_{PRESTO}	97
	E Code source du traitement SIPSFAR	101
	Bibliographie	107
	Table des figures	113
	Table des tableaux	118

Liste des publications

Article de revue

Ait-Allal, D., Weber, R., Dumez-Viou, C., Cognard, I., and Theureau, G. (2012). Rfi Mitigation Implementation For Pulsar Radioastronomy. *Comptes Rendus Physique* 13, 1, pp. 80-87.

Weber, R., Dumez-Viou, Changuel, S., Ait-Allal, D. (2012). Exemples de traitement en temps réel des interférences radioélectriques pour la radioastronomie. *Revue de l'Electricité et de l'Electronique*, n°1, 2012.

Conférences avec proceedings et comité de lecture

Ait-Allal, D., Weber, R., Dumez-Viou, C., Cognard, I., and Theureau, G. (2010). Rfi Mitigation Implementation For Pulsar Radioastronomy. In *Proceedings of EUSIPCO'10 conference*, Aalborg, Danemark.

Ait-Allal, D., Weber, R., Cognard, I., Desvignes, G., and Theureau, G. (2009). Rfi mitigation in the context of pulsar coherent de-dispersion at the nancay radio astronomical observatory. In *Proceedings of EUSIPCO'09 conference*, Glasgow, UK.

Conférences avec proceedings

Changuel, S., Weber, R., Dumez-Viou, C., Ait-Allal, D. (2011) Traitement en temps réel des interférences radioélectriques sur une carte numérique nouvelle génération : UNIBOARD, JS 2011, Les radiotélescopes du futur : Technologies et avancées scientifiques, CNAM, 29 et 30 mars 2011, Paris

Ait-Allal, D., Weber, R., Dumez-Viou, C., Cognard, I., and Theureau, G. (2011) Détection Aveugle de Pulses Géants : Implantation sur GPU, JS 2011, Les radiotélescopes du futur : Technologies et avancées scientifiques, CNAM, 29 et 30 mars 2011, Paris

Ait-Allal, D., Weber, R., Dumez-Viou, C. (2010). RFI mitigation at Nançay Observatory : Impulsive Signal Processing. In *PoS RFI mitigation Workshop 2010*, Groningen, Netherlands

Introduction

Dans les années trente, l'ingénieur Karl Jansky a mis en évidence la présence d'un rayonnement permanent dans la bande décamétrique. Sagittarius A^* , la source radio la plus intense dans le ciel, située au centre de la Voie Lactée, fut alors la première source cosmique à être détectée grâce à son activité électromagnétique dans le domaine radio. Cette découverte fut le point de départ de la radioastronomie.

Dès lors, et plus particulièrement après la seconde guerre mondiale, la radioastronomie se développa véritablement. Depuis, elle est devenue l'un des piliers de l'astronomie moderne. Elle a permis de préciser les connaissances sur les astres déjà connus mais également de découvrir certains des objets les plus intéressants de l'Univers, comme les pulsars, les radiogalaxies ou les quasars.

Le premier pulsar fut découvert en 1967 de façon quelque peu fortuite par Antony Hewish et son étudiante Jocelyn Bell-Burnell avec un instrument mesurant des variations d'un signal de l'émission radio sur des courtes durées (une fraction de seconde). Cet instrument a permis de détecter la variation périodique d'objets qui se sont avérés être des pulsars, le premier d'entre eux porte le nom de PSR B1919+21 (ou CP 1919 à l'époque).

Pour observer et étudier ces radiosources, les radiotélescopes ont dû se développer techniquement. Notamment ils ont dû être équipés de récepteurs spécifiques pour l'observation des pulsars afin de compenser les distorsions (i.e. la dispersion) dues à la traversée du milieu interstellaire. Actuellement, les radioastronomes disposent donc d'instruments très sensibles et observant sur de larges bandes de fréquences.

Or depuis quelques décennies, le domaine des télécommunications a vécu une véritable explosion. Les systèmes de télécommunication se sont multipliés et sont devenus désormais incontournables. Les bandes protégées qui sont exclusivement réservées à la radioastronomie ne sont plus suffisantes pour avoir des observations sans interférence radio fréquence (RFI¹) à cause de leur proximité avec des bandes allouées aux télécommunications. De plus, l'étude des impulsions des pulsars dans certaines bandes qui ne sont pas forcément réservées, permet d'avoir plus d'informations sur leurs caractéristiques d'émission, notamment en basse fréquence [Eilek et al., 2004].

En résumé, le nombre et la taille des bandes de fréquence disponibles pour des observations radioastronomiques de qualité se sont considérablement réduits ces dernières années. Par conséquent, le traitement des RFIs est devenu une préoccupation majeure pour l'exploitation des instruments actuels et dans le développement des radiotélescopes du futur. C'est dans ce contexte que se positionnent les travaux présentés dans cette thèse.

Dans le cadre de l'observation de pulsars connus, l'objectif est de pouvoir statuer en temps réel sur la qualité des données fournies par le radiotélescope et sa chaîne de réception. In fine, l'élimination des données étiquetées comme polluées doit permettre d'améliorer la qualité des mesures faites sur ces pulsars connus.

Dans le cadre de la recherche de nouveaux pulsars ou dans le cadre de la détection d'événements impulsionnels spécifiques, comme par exemple les pulses géants, l'objectif est de proposer une approche qui permette une détection en aveugle (i.e. sans connaissance a priori sur le pulsar) tout en étant robuste aux RFIs.

Le premier chapitre cadre le contexte de travail. Il définit les besoins des scientifiques dans le domaine et décrit les techniques d'observation ; l'approche incohérente utilisée pour la recherche de nouveaux pulsars, et l'approche cohérente utilisée pour la mesure des temps d'arrivée des impulsions quand le pulsar est connu. Finalement, deux modèles du signal d'un pulsar sont donnés, le premier est fondé sur une modélisation temporelle de la distorsion et le second sur une modélisation fréquentielle. Ce dernier modèle sera utilisé dans les simulations présentées dans les chapitres 2 et 3.

Un certain nombre de méthodes de traitement des interférences ont déjà été proposées dans la littérature [Boonstra and Weber, 2009]. Ces méthodes ont été essentiellement orientées pour des observations radioastronomiques classiques, et il existe peu de méthodes qui prennent en compte la spécificité des observations pulsars, sauf pour la méthode zéro DM [Eatough et al., 2009]. Dans le deuxième chapitre, nous présenterons différentes techniques de détection des RFIs en considérant leur utilité dans le contexte des observations de pulsars. Nous

1. c'est l'acronyme en anglais de Radio Frequency Interference et c'est devenu le terme consacré même en français.

analyserons plus en détail les performances de la méthode que nous proposons et qui est fondée sur la cyclostationnarité potentielle des RFIs. Nous proposerons également une première réflexion sur les modalités d'implantation de ce détecteur cyclostationnaire dans une chaîne de traitement en temps réel.

Le troisième chapitre propose une méthode originale, appelée SIPSFAR, qui permet une détection en aveugle et en temps réel des impulsions intenses de pulsar, tout en offrant une certaine robustesse aux RFI. Elle est basée sur la transformée de Fourier à deux dimensions et la transformée de Radon. Une étude théorique comparative avec la méthode classique de recherche de d'impulsions individuelles ainsi que des tests sur des données réelles sont proposés.

Le dernier chapitre présente les résultats obtenus avec SIPSFAR sur un grand volume de données réelles enregistrées lors d'un grand relevé du ciel, couvrant le plan Galactique Nord. L'implantation de SIPSFAR a été réalisée sur un GPU (Graphic Unit Processor). Finalement, une analyse statistique a été faite en vue de comparer les performances de la méthode classique et de la nouvelle méthode que nous avons proposée.

Chapitre 1

L'observation radio des pulsars

1.1 Introduction

Les pulsars font partie des nombreux objets célestes que les astronomes observent en ondes radio ([Lyne and Graham-Smith, 2006]). Un pulsar est une étoile à neutrons fortement magnétisée en rotation rapide. Elle est le reste de l'explosion, appelée supernova, d'une étoile massive pesant plus de huit fois la masse du soleil environ, laissant derrière elle un cœur très compact essentiellement constitué de neutrons. En effet, avec environ 1.4 masses solaires pour un diamètre d'une vingtaine de kilomètres, la densité obtenue est extrême et seuls les neutrons la supportent. Par conservation du moment cinétique, l'étoile accélère sa rotation pour atteindre quelques dizaines de tours par seconde. La conjonction de la rotation très rapide et du très fort champ magnétique génère deux faisceaux radio qui s'échappent de part et d'autre au niveau des pôles magnétiques de l'étoile à neutrons, balayant l'espace régulièrement comme un phare au bord de la mer (voir figure 1.1). L'étoile à neutrons magnétisée alors observée en ondes radio est appelée un pulsar. Plus de 2000 pulsars¹ ont été trouvés depuis la découverte initiale faite par les astrophysiciens britanniques Jocelyn Bell Burnell et Antony Hewish à Cambridge en 1967 [Hewish et al., 1968].

Il est très important pour les astronomes d'étudier la forme des impulsions des pulsars et de mesurer leurs temps d'arrivée, car cela permet de comprendre et vérifier la physique des étoiles à neutrons, la relativité générale et d'autres sujets prédits par les théories de l'évolution stellaire. Cependant, les signaux de pulsars traversent le milieu interstellaire (ISM) avant d'arriver jusqu'à nous. La composante ionisée de ce milieu provoque principalement la dispersion des impulsions. Concrètement, les impulsions observées dans les hautes fréquences arrivent plus tôt au radiotélescope que celles observées dans les basses fréquences. Les niveaux de flux radio reçus sont si faibles que les impulsions sont en général à peine détectables individuellement. Notons aussi que, même après dédispersion, le niveau de flux radio n'est pas stable d'une impulsion à l'autre. De nombreuses études nécessitent l'intégration d'un grand nombre d'impulsions successives afin d'augmenter le rapport signal sur bruit.

Alors que la majorité des pulsars ont été découverts par une analyse de Fourier et une signature périodique, certains pulsars ont été découverts grâce à une propriété atypique de leurs émissions radio, à savoir le phénomène des impulsions individuelles géantes. L'impulsion dite géante est beaucoup plus intense que la moyenne (>100 - 1000 fois) et donc détectable individuellement malgré sa durée extrêmement courte (une fraction de microseconde parfois). Un autre type d'étoiles à neutrons magnétisées, non détecté par une analyse de Fourier, produit épisodiquement des impulsions individuelles ordinaires et détectables. Elles ont toutes les caractéristiques des impulsions individuelles que l'on parvient à observer sur les pulsars les plus intenses, si ce n'est qu'elles sont très rares. Ce sont les RRATs (Rotating Radio Transients) dont nous ne recevons qu'une impulsion de temps à autre, ce qui rend d'ailleurs très difficile la détermination exacte de leurs paramètres de rotation. Ces deux phénomènes font donc l'objet d'une procédure de détection particulière, nommée détection d'impulsion individuelle, complémentaire à l'analyse de Fourier.

Pour faire face aux problèmes rencontrés lors des observations pulsars, les astronomes et les ingénieurs ont mis au point une architecture instrumentale particulière et des algorithmes de traitement du signal afin d'inverser l'effet de l'ISM. Dans la première section, on s'intéressera à l'architecture employée à la station de radioastronomie de Nançay. Dans la deuxième section, on présentera un signal synthétique modélisant un signal pulsar. Finalement, l'impact des interférences radio fréquence (RFI) sur la qualité des observations sera abordé.

1. ATNF Pulsar Catalog <http://www.atnf.csiro.au/research/pulsar/psrcat/>

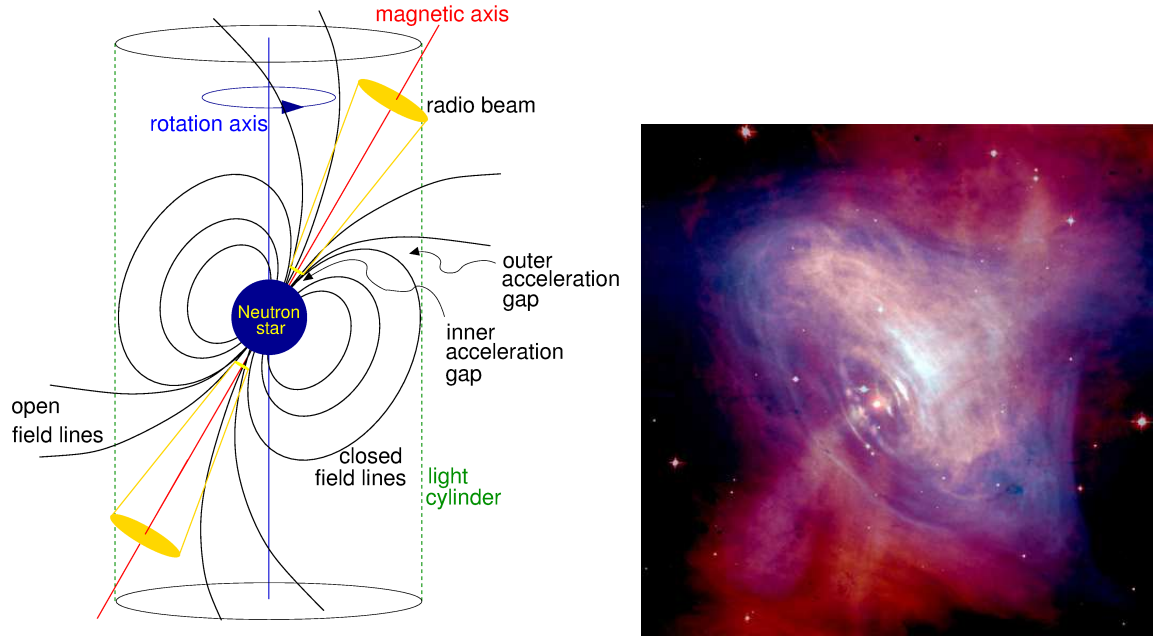


FIGURE 1.1 – (gauche) Modèle de l'étoile à neutrons et de sa magnétosphère avec l'émission radio s'échappant au niveau de l'axe magnétique ([Lorimer and Kramer, 2005]). (droite) Environnement immédiat du pulsar du Crabe (PSR B0531+21), en optique (image Hubble, en rouge) et en X (image Chandra, en bleu).

1.2 L'instrumentation pour l'observation des pulsars

1.2.1 Le radiotélescope de Nançay

Le grand radiotélescope décimétrique de Nançay fut inauguré en 1965 et il a été construit principalement pour l'étude des comètes, des galaxies, des enveloppes d'étoiles et des pulsars, initialement autour de la raie de l'hydrogène neutre à 1.4GHz. Il faut pour observer le signal radio de ces sources très lointaines souvent très faibles une grande surface collectrice. Cet instrument est symétrique par rapport à un plan Nord-Sud et un astre ne peut être observé que lorsqu'il passe au voisinage de ce plan, appelé plan méridien. Les ondes reçues sont réfléchies par un premier miroir (figure 1.2) qui est un plan de 200 m x 40 m, et dont on peut modifier l'inclinaison pour aller observer des sources plus ou moins hautes dans le ciel. Les ondes arrivent ensuite sur un miroir fixe, situé à 460 m du premier, et mesurant 300 m x 35 m. Celui-ci est sphérique, de 560 m de rayon. Il concentre les ondes en un point, son foyer, situé entre les deux miroirs, où les ondes sont recueillies par des récepteurs placés derrière des cornets. A cause de la rotation de la Terre, la zone de ciel visée se déplace lentement, d'Est en Ouest. Le foyer où se concentrent les ondes se déplace donc sur une "surface focale", et les récepteurs doivent suivre ce mouvement.

Le radiotélescope décimétrique (ou NRT pour Nançay Radio Telescope) est l'instrument le plus imposant de l'Observatoire de Nançay. Avec sa surface collectrice effective de 7000 m², il reste le quatrième plus grand du monde dans cette gamme de fréquence. Le programme de chronométrie des pulsars a débuté fin 1988 avec le suivi de seulement deux objets, PSR B1937+21 et PSR B1821-24 principalement pour l'étude du milieu interstellaire à l'époque. Le radiotélescope fut rénové profondément en 1998-2000. Tout l'ensemble chariot-cornet-récepteur fut remplacé et un gain en sensibilité d'un facteur 2.2 obtenu. Il est aussi désormais moins sensible au rayonnement parasite et intègre deux nouveaux récepteurs qui permettent respectivement une couverture de 1.1 GHz à 1.8 GHz et de 1.7 GHz à 3.5 GHz avec une bande passante maximum beaucoup plus large de ~500 MHz avec une température de système² T_{sys} de 35K au lieu de 50K pour l'ancien système.

1.2.2 La dispersion par le milieu interstellaire

Le milieu interstellaire, constitué d'un gaz ténu et ionisé avec une densité électronique $n_e \approx 0.03 \text{ cm}^{-3}$, introduit un retard dispersif dans la propagation des ondes radio qui est fonction de leur fréquence. Ce phénomène

2. où le bruit du système correspond à celui d'une résistance à la même température



FIGURE 1.2 – Le radiotélescope décimétrique de Nançay

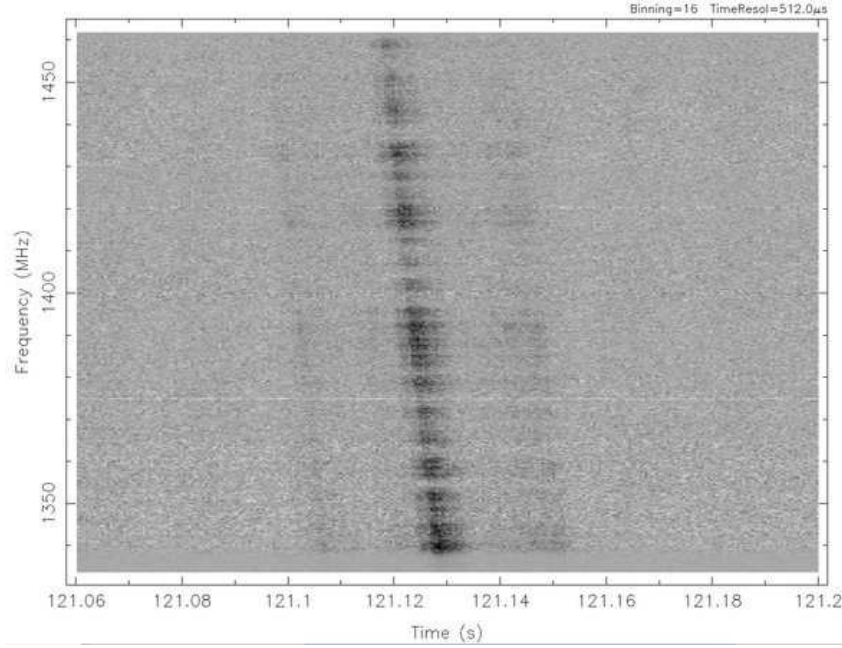


FIGURE 1.3 – Dispersion d'une impulsion individuelle du pulsar PSR B0329+54 observée à Nançay en avril 2009 sur une bande totale de 128MHz centrée à 1398MHz.

de dispersion par le milieu interstellaire fut identifié dès la découverte du premier pulsar. Un exemple de signal pulsar dispersé est donné en figure 1.3 avec une impulsion individuelle du pulsar lent et intense PSR B0329+54 observée sur une bande totale de 128MHz centrée à 1398MHz. Par rapport à une onde électromagnétique de fréquence infinie, le retard observé à la fréquence f s'écrit :

$$t = \frac{kDM}{f^2} \quad (1.1)$$

où $k = 4.148808(3) \times 10^3 \text{ cm}^3 \text{ pc}^{-1} \text{ MHz}$ est la constante de dispersion, et DM la mesure de dispersion comme étant l'intégrale sur la ligne de visée Terre - pulsar de la densité d'électrons $n_e(x)$:

$$DM = \int_0^d n_e(x) dx \quad (1.2)$$

Le délai différentiel δt en secondes, sur une bande de fréquence $\delta f = f_2 - f_1$ centrée sur $f = \frac{f_1 + f_2}{2}$, est alors, en supposant $\delta f \ll f$:

$$\delta t \approx 2 \times k \times \frac{\delta f}{f^3} \times DM \quad (1.3)$$

Étant donné que les impulsions des pulsars sont étroites et rapprochées, la dispersion devient problématique quand on observe sur une large bande passante, à basse fréquence ou des objets avec de grands DMs. Dans le plan galactique, on trouve des pulsars avec des valeurs de DM comprises entre 2 et $\sim 1500 \text{ pc.cm}^{-3}$. A titre

d'exemple, pour le pulsar B1937+21 (un des pulsars millisecondes les plus rapides avec une période de 1.55 ms et une mesure de dispersion DM de 71 pc.cm⁻³), la dispersion dans une bande de 10 MHz centrée à 1.4 GHz est de $\delta t = 2.15$ ms. Sur cette bande de fréquence relativement petite, l'impulsion est ainsi déjà élargie sur plus d'une période de rotation, la rendant donc indétectable si le signal est intégré sur toute la bande sans précaution ou compensation particulière. Pour combattre les effets de la dispersion, il existe deux principales méthodes :

- La méthode dite incohérente (ou par banc de filtres).
- La dédispersion cohérente.

Ces méthodes seront abordées dans les sections suivantes.

1.2.3 La dédispersion incohérente

La méthode la plus simple pour compenser la dispersion est de découper la bande de fréquence d'entrée en un certain nombre de canaux indépendants. Ensuite, on applique le retard adéquat pour chaque canal, pour que les portions de l'impulsion arrivent en même temps en sortie. Le retard approprié doit être calculé en fonction de la mesure de dispersion (contenu électronique intégré le long de la ligne de visée) DM et de la fréquence centrale de chaque canal f_0 . L'équation 1.1 permet de déterminer ce retard t_{f_0} en milliseconde par rapport à une fréquence de référence f_{ref} :

$$t_{f_0} = 4.15 \times 10^6 (f_{ref}^{-2} - f_0^{-2}) DM \quad (1.4)$$

Dans l'équation (1.4) les fréquences sont en MHz et la mesure de dédispersion DM est en pc.cm⁻³ (voir [Lorimer and Kramer, 2005], chapitre 5 et 6). Soit S_{jj} la puissance du $j^{ème}$ échantillon temporel dans le $l^{ème}$ canal de fréquence. Pour L_{chans} canaux, le $j^{ème}$ échantillon temporel des données dédispersées D_j est :

$$D_j = \sum_{L_{chans}}^{l=1} S_{j+k(l),l} \quad (1.5)$$

où $k(l)$ est l'entier le plus proche correspondant au retard induit (Eq. 1.4) par la dispersion dans le canal numéro l .

1.2.4 La dédispersion cohérente

La dispersion par le milieu interstellaire peut également être complètement enlevée par ce qu'on appelle la dédispersion cohérente qui a été proposée par Hankins et Rickett (1975). En effet, cette méthode tire parti du fait que la dispersion intervient uniquement sur la phase du signal et peut être comparée à l'effet d'un filtre dont la fonction de transfert est H (Equ. 1.7).

La relation entre le signal original $v(t)$ et le signal reçu dispersé $v_r(t)$ devient très simple dans le domaine des fréquences. Pour un signal de fréquence f , centré sur une fréquence f_0 et avec une bande passante limitée Δf , on peut écrire :

$$V_r(f_0 + f) = V(f_0 + f)H(f_0 + f) \quad (1.6)$$

$V(f)$ et $V_r(f)$ sont les transformées de Fourier respectives du signal émis $v(t)$ et du signal reçu $v_r(t)$. La fonction de transfert H s'écrit :

$$H(f_0 + f) = \exp\left(+j \frac{2\pi DM f^2}{2,41 \cdot 10^{-10} (f + f_0) f_0^2}\right) \quad (1.7)$$

où DM est toujours la mesure de la dispersion directement observée.

La dédispersion cohérente consiste donc à appliquer l'inverse de cette fonction de transfert sur le signal dispersé reçu. Ce traitement est effectué dans le domaine fréquentiel, en utilisant les transformées de Fourier, ainsi que la multiplication par la fonction de transfert d'un autre filtre pour minimiser l'aliasing. La combinaison de l'inverse de la fonction de transfert de l'ISM et du filtre antialiasing est nommée la fonction Chirp $C(f)$. Ainsi, on a :

$$V(f_0 + f) = V_r(f_0 + f)C(f_0 + f) \quad (1.8)$$

La dédispersion cohérente est maintenant couramment utilisée pour effectuer des mesures précises de temps d'arrivée de pulsars rapides et ultra-stables. En effet, la dédispersion cohérente permet de retrouver la forme de

l'impulsion originale sans aucun élargissement dû à la largeur des canaux de fréquences utilisés, contrairement aux instrumentations à dédispersion incohérente. Bien entendu, cette technique ne peut s'appliquer que lorsque le pulsar est connu d'avance, et lorsque la valeur de la mesure de dispersion DM est bien mesurée.

1.2.5 L'algorithme de l'empilement 'Folding'

Les impulsions individuelles des pulsars sont connues pour être variables. De plus, ces objets sont des sources radio très faibles. Seul l'empilement de milliers d'impulsions individuelles permet d'avoir un profil moyen caractéristique et stable. Ceci permet de réduire le bruit sur le profil de l'impulsion et ainsi minimiser les erreurs d'incertitude sur l'estimation du temps d'arrivée. Chaque échantillon étant précisément daté, la phase rotationnelle du pulsar peut être obtenue à l'aide des paramètres connus du pulsar (période, dérivée de la période, position, paramètres orbitaux en cas de pulsar binaire orbitant avec une étoile compagnon, etc...). Ceci se fait avec une éphéméride calculée par le logiciel TEMPO³. L'algorithme du Folding peut être résumé comme suit :

1. Créer un tableau pour la sauvegarde du profil empilé avec N_{pts} cellules, identiquement espacées dans la période du pulsar, en général $N_{pts} = 2048$.
2. Calculer la phase par rapport à la période du pulsar grâce à TEMPO pour chaque échantillon de la série temporelle dédispersée.
3. Additionner cet échantillon dans la case correspondante du tableau créé.
4. Normaliser le tableau par le nombre d'éléments additionnés dans chaque case du tableau.

1.2.6 L'architecture et les ressources

Les impulsions périodiques reçues des pulsars peuvent être mesurées et enregistrées avec des radiotélescopes et des instrumentations dédiés. La figure 1.4 décrit les étapes typiques d'une telle architecture. Si le pulsar est déjà connu, les radioastronomes sont intéressés par l'analyse précise du signal, par exemple pour la mesure précise du temps d'arrivée de l'impulsion (c'est la chronométrie, ou chronométrage selon les auteurs), et un traitement de dédispersion cohérente ([Lorimer and Kramer, 2005]) est effectué alors en temps réel. S'il s'agit de chercher un nouveau pulsar, l'instrumentation sera configurée pour fournir un grand nombre de canaux de fréquence étroits sur lesquels de nombreux essais de dédispersion incohérente par retard temporel seront effectués en temps différé pour ensuite chercher un signal périodique par analyse de Fourier.

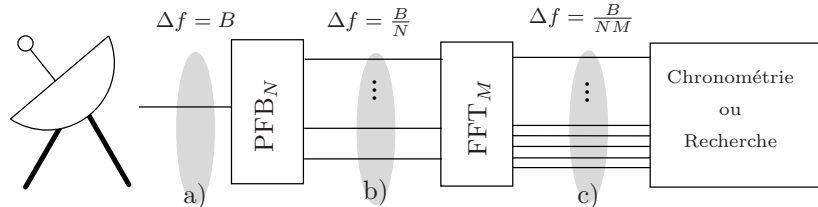


FIGURE 1.4 – L'architecture typique d'une instrumentation 'pulsar' : a) étape 1 : acquisition de la forme d'onde b) étape 2 : banc de filtres polyphasés (PFB) pour séparer la bande de fréquence d'entrée (habituellement des centaines de MHz) en petits morceaux ($B/N \sim$ quelques MHz) c) étape 3 : après transformée de Fourier sur quelques MHz, soit un filtre approprié est appliqué pour dédisperser avant de revenir en temps pour effectuer la mesure de temps d'arrivée, soit une transformée de Fourier inverse particulière produit des séries temporelles sur de très petits canaux ($B/NM \sim 0.25$ MHz) qui seront écrites sur disque en vue d'une exploration ultérieure en DM et en périodicité pour rechercher de nouveaux pulsars.

Comme dans beaucoup de radiotélescopes, les ondes électromagnétiques sont captées à Nançay par un récepteur refroidi sous forme de deux polarisations linéaires orthogonales. Ensuite, le transport du signal des fréquences hautes dites 'ciel' f_0 vers des fréquences plus basses, plus facile à analyser, se fait selon une chaîne hétérodyne classique où alternent filtres passe-bande et oscillateurs locaux.

Depuis juillet 2008, la chronométrie ou la recherche (Timing ou Survey) se fait dans un système innovant développé spécialement à la Station de Nançay en collaboration avec P. Demorest et D.C. Backer de UC Berkeley (voir [Demorest, 2007]). Après un dernier décalage pour ramener les signaux complexes d'une bande passante totale $\Delta f = 128$ MHz autour de la fréquence zéro où se fait la conversion analogique-numérique, les données

3. <http://www.atnf.csiro.au/people/pulsar/tempo/>

numérisées sur 8 bits sont transmises au filtre polyphase d'un processeur programmable (FPGA). Ce filtre sert à découper la bande totale en sous bandes de 4MHz, plus facile à traiter ultérieurement. Le système contient quatre ordinateurs à microprocesseurs qui lisent les données sorties du FPGA et les ré-arrangent pour construire des séries temporelles chronologiques de chaque canal de fréquence de 4MHz. Les données sont ensuite envoyées à quatre GPUs (Graphic Processing Unit - NVIDIA 88 GTX - 128 parallel processors) pour la partie finale du traitement constituée de plusieurs algorithmes visant à dédispenser le signal du pulsar, construire les polarisations croisées (en vu de l'obtention des paramètres de Stokes), puis empiler les impulsions périodiques les unes sur les autres. A interval de temps régulier est ainsi stocké sur disque une impulsion moyenne avec tous les paramètres permettant son utilisation ultérieure (fréquence, temps). Dès que les impulsions sont intégrées suffisamment longtemps, le profil obtenu peut être vu comme une sorte d'empreinte unique.

1.3 Modèle du signal d'un pulsar

Pour être en mesure de tester différents détecteurs de parasites dans le cas spécial des observations pulsars, en présence donc d'impulsions dispersées individuelles, nous allons construire une modélisation 'forme d'onde' du signal du pulsar. Il s'agit de produire une série temporelle représentant l'évolution de la **tension** mesurée par un radiotélescope qui est directement proportionnelle au champ électromagnétique de l'onde radio reçue par son détecteur. Cette série temporelle représentera le signal reçu sur une bande de fréquence donnée et pendant un temps donné. Elle pourra présenter différents niveaux de bruit et sera donnée en entrée des différents détecteurs testés dans le prochain chapitre. En effet, un signal synthétique modélisant l'impulsion du pulsar demeure essentiel afin de valider rapidement les différents algorithmes avant de les tester sur des données réelles qui sont plus longues à exploiter.

Rappelons d'abord que certaines propriétés des émissions radio des pulsars, comme leurs périodes et leurs formes moyennes sont stables. En revanche, les propriétés des impulsions individuelles changent rapidement et d'une manière aléatoire, notamment l'intensité et la forme. Donnons d'abord quelques définitions avant d'aller plus loin dans cette section.

On définit une impulsion comme étant la forme de l'émission radio reçue lorsque le faisceau du pulsar est dirigé vers le radiotélescope. Du fait de la stabilité de la rotation de l'étoile à neutrons magnétisée qu'est le pulsar, les impulsions reviennent régulièrement avec une périodicité constante propre à chaque pulsar. L'impulsion présente un seul maximum le plus souvent, mais peut aussi présenter deux maxima à environ une demi-période l'un de l'autre (et il est supposé que les deux pôles magnétiques "éclairaient" alors la Terre) ou même des formes quelconques et très larges ([Lorimer and Kramer, 2005]). Toutefois, nous ne nous intéresserons pas à cet aspect car ce n'est pas l'objet d'étude de ce mémoire.

Le profil moyen, le profil intégré, ou simplement le profil du pulsar est la moyenne synchrone de plusieurs impulsions successives au long de sa période, dont la construction a été décrite en section 1.2.5.

L'impulsion individuelle du pulsar peut être considérée comme un signal gaussien modulé en amplitude [Rickett, 1975]. L'enveloppe de l'amplitude peut être considérée comme une empreinte du pulsar dans une bande de fréquence donnée : c'est le profil intégré. Même si le profil réel peut être complexe, ici pour simplifier, une forme gaussienne $f(t)$ comme enveloppe de l'impulsion synthétique a été choisie. Elle sera multipliée par un signal $b(t)$ qui est un bruit blanc gaussien de moyenne nulle et de variance σ_b^2 .

$$v(t) = b(t) \times f(t) \quad (1.9)$$

où la gaussienne sera $f(t) = e^{-\frac{1}{2}(\alpha \frac{n}{N/2})^2}$ avec la variable temps t représentée par un entier n satisfaisant $-\frac{N}{2} \leq n \leq \frac{N}{2}$ et avec $\alpha \geq 2$ est un facteur proportionnel à l'inverse de l'écart type. La longueur de l'impulsion synthétique est $l_p = N + 1$.

Nous allons aborder deux approches différentes pour produire la forme d'onde d'une impulsion dispersée. Une première approche dite fréquentielle est identique à la dispersion cohérente. Elle repose sur la modélisation d'un retard et la construction d'une fonction de transfert appliquée dans le domaine de Fourier. La seconde approche est dite temporelle et est abordée pour compléter la présentation mais ne sera pas utilisée par la suite.

Modélisation de la dispersion dans le domaine fréquentiel

Le retard en fonction de la fréquence est ici modélisé de façon très simple puis appliqué au signal dans l'espace de Fourier pour disperser comme le milieu interstellaire. Enfin une transformée de Fourier inverse permettra de retrouver le signal "forme d'onde" temporel.

Le retard est modélisé de façon particulière pour obtenir une forme parabolique qu'un seul paramètre de courbure permet de contrôler. La forme exacte du retard choisi ci-dessous n'est pas physique car nous avons simplifié le retard normalement inversement proportionnel au carré de la fréquence par une parabole. Mais cette approche permet de n'avoir qu'un seul paramètre de courbure sans se soucier de la mesure de dispersion DM et de la fréquence radio d'observation. Cela ne change évidemment pas la nature des simulations du chapitre suivant et les conclusions obtenues. Écrivons donc le retard comme :

$$\tau(f) = af^2 + bf + c \quad (1.10)$$

Le plan temps-fréquence du signal $x(t)$ est constitué de N échantillons temporels et l'axe des fréquences est normalisé entre 0 et 1. Pour que l'impulsion une fois dispersée occupe tout le plan temps-fréquence, les retards aux bornes de fréquence $f = 0$ et $f = 1$ sont $\tau(f = 0) = N$ et $\tau(f = 1) = 0$. Ceci implique :

$$\begin{aligned} c &= N \\ a + b + c &= 0 \end{aligned} \quad (1.11)$$

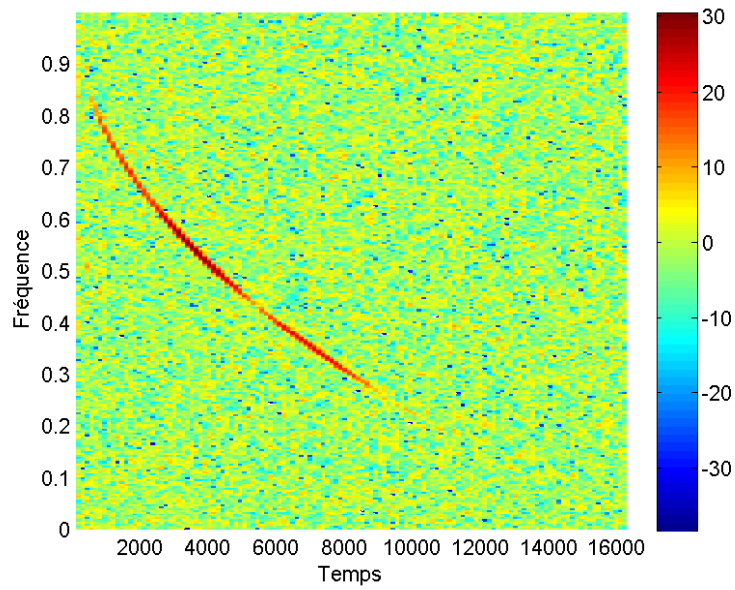


FIGURE 1.5 – Plan temps fréquence d'une impulsion individuelle modélisée dans le domaine fréquentiel par l'équation 1.17. L'axe de fréquence est normalisé entre 0 et 1, alors que l'axe du temps l'est entre 0 et le nombre de points N .

La dérivée du retard en fonction de f est $\frac{\delta\tau}{\delta f} = 2af + b$, avec d_0 et d_1 les valeurs des dérivées, on a alors :

$$\begin{aligned} \frac{\delta\tau}{\delta f}|_{f=0} &= d_0 = b \\ \frac{\delta\tau}{\delta f}|_{f=1} &= d_1 = 2a + d_0 \end{aligned} \quad (1.12)$$

L'équation 1.11 et l'équation 1.12 donnent :

$$a = \frac{d_1 - d_0}{2} \quad (1.13)$$

et :

$$\frac{d_1 - d_0}{2} + d_0 + N = 0 \quad (1.14)$$

Avec les équations 1.13 et 1.14 on a :

$$a = -d_0 - N \quad (1.15)$$

Et donc on a $d_1 = -d_0 - 2N$ avec $-N \leq d_1 \leq 0$ les limites de la courbure de l'impulsion. On en déduit alors : $-2N \leq d_0 \leq -N$. On pose $d_0 = -N - \alpha N$ avec $\alpha \in [0, 1]$, 0 pour une dispersion linéaire et 1 pour une dispersion quadratique.

Rappelons que l'application d'un retard τ à un signal de type $e^{j2\pi ft}$ permet d'obtenir un signal de la forme $e^{j2\pi f(t-\tau)} = e^{j2\pi ft} \times e^{-j2\pi f\tau}$ où $e^{-j2\pi f\tau}$ est la fonction de transfert de l'opérateur "retard". La phase de cette fonction de transfert est $\phi(f) = -j2\pi f\tau$ dont la dérivée par rapport à la fréquence f est proportionnelle à τ . Ainsi, si τ est, à un facteur près, la dérivée par rapport à f de la phase $\phi(f)$ de la fonction de transfert $e^{j\phi(f)}$, alors à partir d'une fonction retard $\tau(f)$ quelconque, on retrouve $\phi(f)$ par intégration selon f . Ici, on a défini une fonction retard $\tau(f)$ à notre convenance et on va pouvoir en déduire par intégration notre fonction de transfert.

La fonction qui modélise alors la fonction de transfert $H(f)$ est $e^{-j\varphi(f)}$, avec $\varphi(f)$ qui est l'intégrale du retard $\tau(f)$:

$$H(f) = e^{-j2\pi(\frac{a}{3}f^3 + \frac{b}{2}f^2 + cf)} e^{-j2\pi d} \quad (1.16)$$

avec $2\pi d$ qui est une phase constante aléatoire.

Finalement, comme le signal $v(t)$ doit être dispersé dans le domaine fréquentiel avec le filtre $H(f)$ calculé plus haut (équation 1.16), le signal dispersé est retrouvé avec une transformée de Fourier inverse pour revenir dans le domaine temporel :

$$p(t) = TF^{-1}[TF(v(t)) \times H(f)] \quad (1.17)$$

La figure 1.5 montre une série temporelle forme d'onde synthétisée sous forme de spectre dynamique (plan temps-fréquence où l'évolution du contenu spectral de la bande de fréquence est tracée en fonction du temps). Notons que c'est au niveau de $H(f)$ qu'il est possible de multiplier par une fonction qui permet de simuler la scintillation interstellaire en faisant en sorte que l'amplitude du pulsar soit variable selon la fréquence radio.

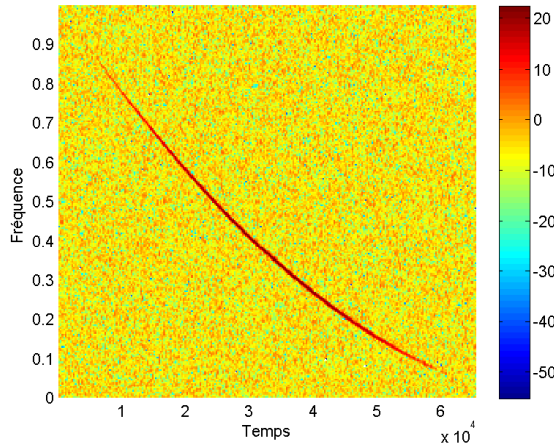


FIGURE 1.6 – Plan temps fréquence d'une impulsion individuelle modélisée dans le domaine temporel par l'équation 1.20. L'axe de fréquence est normalisé entre 0 et 1, alors que l'axe du temps l'est entre 0 et le nombre de points N .

Modélisation de la dispersion dans le domaine temporel

Nous allons ici d'abord considérer l'impulsion comme une fonction de Dirac à laquelle nous allons ajouter un retard de type dispersif. Pour obtenir une impulsion de forme raisonnable, il suffirait d'accoler un certain nombre de fonctions Dirac les unes à côté des autres. La dispersion peut être vue comme une fonction 'chirp' où la fréquence change au cours du temps.

Pour un signal dans une bande de fréquence Δf et ayant un retard de dispersion total t_{DM} , la fréquence peut être écrite comme suit :

$$f(t) = \frac{\alpha \Delta f}{t_{DM}^2} t^2 - \frac{\Delta f(\alpha + 1)}{t_{DM}} t + \frac{\Delta f}{2} \quad (1.18)$$

avec $t \in [0, t_{DM}]$ et α est un facteur caractérisant le degré de courbure (la mesure de dispersion DM). Donc la phase en fonction du temps peut s'écrire :

$$\varphi(t) = 2\pi \left(\frac{\alpha \Delta f}{3t_{DM}^2} t^3 - \frac{\Delta f(\alpha + 1)}{2t_{DM}} t^2 + \frac{\Delta f}{2} t + d \right) \quad (1.19)$$

Avec $v(t)$ une fonction qui permet de moduler l'intensité du pulsar pour modéliser la scintillation, l'impulsion dispersée est obtenue avec :

$$p(t) = v(t)e^{j\varphi(t)} \quad (1.20)$$

Cette approche par le domaine temporel n'a pas été utilisée par la suite.

1.4 L'impact des RFIs sur la qualité des observations

Après la détection en 1932 par Karl G. Jansky de l'émission radio de notre galaxie, les découvertes des émissions radio de plusieurs sources cosmiques se succédèrent à la suite, et la communauté scientifique réclama alors un statut officiel auprès de l'ITU (International Telecommunication Union) afin que la radioastronomie soit considérée comme un service à part entière de la bande radio-fréquence. Au fur et à mesure de nouvelles bandes furent négociées et différents statuts ont été obtenus. Dans quelques bandes étroites, un statut d'utilisateur primaire pour lequel l'utilisation est pleinement reconnue a été donné à la radioastronomie. Un utilisateur secondaire, lui, ne devra pas produire d'interférence susceptible de perturber un utilisateur primaire et ne pourra exiger protection des perturbations causées par un utilisateur primaire de la même bande. Un extrait du tableau d'allocation des fréquences en 2004 est donné par le tableau 1.1.

Allocation ITU	Application	Statut
13.36 - 13.41 MHz		Primaire
25.55 - 25.67 MHz		Primaire
37.50 - 38.25 MHz	Continuum	Secondaire
150.05 - 153 MHz	Continuum, Pulsar, Solar	Secondaire
322 - 328.65 MHz	Continuum, VLBI	Primaire
406.1 - 410 MHz	Continuum, Pulsar	Primaire
608 - 614 MHz	Continuum, VLBI	Secondaire
1300 - 1400 MHz	Raies spectrales	Secondaire
1400 - 1427 MHz	Raies spectrales	Primaire
1610.6 - 1613.8 MHz	Raies spectrales	Primaire
1660 - 1660.5 MHz	Continuum, VLBI	Secondaire
1660.5 - 1670 MHz	Continuum, VLBI	Primaire
2200 - 2290 MHz	VLBI	Secondaire
2665 - 2690 MHz	Continuum	Secondaire
2690 - 2700 MHz		Primaire
4800 - 5000 MHz	Continuum, VLBI	Secondaire
5000 - 5030 MHz	VLBI	Primaire
8215 - 8400 MHz	VLBI	Secondaire

TABLE 1.1 – Extrait du tableau d'allocation des fréquences, ITU, Mai 2004

Malheureusement la radioastronomie n'est pas toujours considérée comme un utilisateur primaire dans beaucoup de ces bandes. De plus, certaines bandes autrefois considérées comme peu intéressantes sont maintenant très utiles scientifiquement. Les observations doivent donc parfois se faire en présence de puissants émetteurs radio artificiels. Nous savons maintenant que les émissions radio des pulsars sont plus intenses en basse fréquence (ici ~ 150 MHz). Mais depuis la découverte du premier pulsar, les observations ont été majoritairement faites dans des bandes de fréquences plus hautes (300-2000 MHz) pour limiter les altérations provoquées par le milieu interstellaire sur les signaux impulsionnels, par le bruit dû à la température élevée du ciel et par les

effets de l'ionosphère dans les basses fréquences. Malgré ces obstacles, scientifiquement il est très important d'observer les pulsars en basse fréquence. Les quelques études qui ont été faites ont montré l'importance de ces observations pour l'étude de la morphologie des impulsions et leurs distributions d'énergie [Eilek et al., 2004]. Grâce aux nouvelles générations de radiotélescopes et de récepteurs, l'observation des pulsars en basse fréquence est désormais possible (LOFAR : LOw Frequency ARray⁴). Sauf que dans ces bandes de fréquence, la radioastronomie n'est qu'utilisateur secondaire, ces bandes sont donc fortement occupées par les télécommunications et ceci altère la qualité des observations d'une manière considérable. La présence de ces brouilleurs détériore le rapport signal sur bruit du profil intégré du pulsar et détériore la mesure précise des temps d'arrivée. De plus elle augmente la probabilité de fausse détection quand il s'agit de la recherche de nouveaux pulsars ou de la détection d'événements impulsionnels intermittents. Ainsi, dans certains cas on peut perdre jusqu'à 50% des données. Il faut donc implémenter des algorithmes robustes de traitement d'interférences dans la chaîne de traitement temps réel pulsar.

Dans le chapitre suivant, nous allons présenter plus en détail la problématique du traitement des interférences dans le contexte des observations des pulsars et nous allons faire l'étude de plusieurs méthodes de détection ainsi qu'une étude comparative pour conclure quelles méthodes sont plus adaptées au contexte de l'étude.

4. <http://www.lofar.org/>

Chapitre 2

Traitement des interférences pour les observations de pulsars

2.1 Introduction

Dans ce chapitre, après une introduction générale sur le traitement des RFI (section 2.2) qui nous permettra de poser plus en détail notre problématique, nous présenterons plusieurs techniques actuelles de détection des RFI fondées sur des mesures de puissance (section 2.3). En effet, l'utilisation d'un critère de puissance est une approche très majoritaire dans les techniques de dépouillement et de nettoyage de données en radioastronomie. La raison principale en est que la mesure de puissance est le matériau de base délivré au radioastronome par un radiotélescope.

Or cette approche peut être limitative, notamment dans le cas qui nous concerne, c'est à dire l'observation des pulsars. Par exemple, par le seul critère de puissance, comment distinguer automatiquement l'impulsion pulsar, notamment géante, d'un quelconque sursaut d'une émission radar ?

Dans ce chapitre, nous proposons donc une autre approche basée sur les propriétés cyclostationnaires des signaux de télécommunication. Les détails théoriques de ce critère seront présentés à la section 2.4 puis ses performances seront analysées au travers de simulations sur des signaux synthétiques (section 2.6). L'objectif sera alors, dans un premier temps, d'évaluer la robustesse du critère cyclostationnaire en présence d'un pulse de pulsar, et dans un deuxième temps, d'étudier la sensibilité du dit détecteur en présence d'une RFI. A titre de comparaison, cette étude sera également réalisée sur un critère fondé sur la non-gaussianité de l'interférence (section 2.5). Conceptuellement, ce critère, proposé par Nita et al. [Nita et al., 2007], peut être une alternative intéressante au problème de discrimination entre sources cosmiques (par nature gaussiennes) et RFI (par construction, rarement gaussiennes).

A terme, l'objectif étant d'implanter ces techniques sur la chaîne de traitement en temps réel d'un radiotélescope, nous avons étudié l'impact des résolutions fréquentielles et temporelles sur les performances des deux critères de détection précédents (i.e. cyclostationnarité et non-gaussianité). Cette étude (section 2.7) s'est faite dans le cadre du projet européen UNIBOARD (section 2.8).

Enfin, pour clore ce chapitre, nous proposerons des simulations sur des données réelles issues des radiotélescopes LOFAR et UTR2 (section 2.9).

2.2 Introduction générale sur les interférences radioélectriques

Le traitement des interférences pour les observations radioastronomiques peut s'avérer une tâche délicate. Il vise à trouver le juste milieu entre l'élimination des interférences et la préservation d'un minimum de données nécessaire à la détection des sources cosmiques, souvent de faible intensité. D'autant plus qu'il faut savoir qu'aucune technique ne peut être considérée comme LA solution pour toutes les applications en radioastronomie [Boonstra et al., 2010].

En effet, selon le type d'instrument, d'observation et d'interférences, trois scénarios sont possibles en fonction du degré du mélange entre le signal cosmique et le signal d'interférence [Weber et al., 2004].

- Mélange total : Il se produit quand aucune discrimination n'est possible entre le signal radio cosmique et l'interférence. C'est généralement le cas quand les émetteurs diffusent en continu sur le même domaine temps-fréquence que la source cosmique. Deux approches peuvent être alors considérées : le filtrage spatial

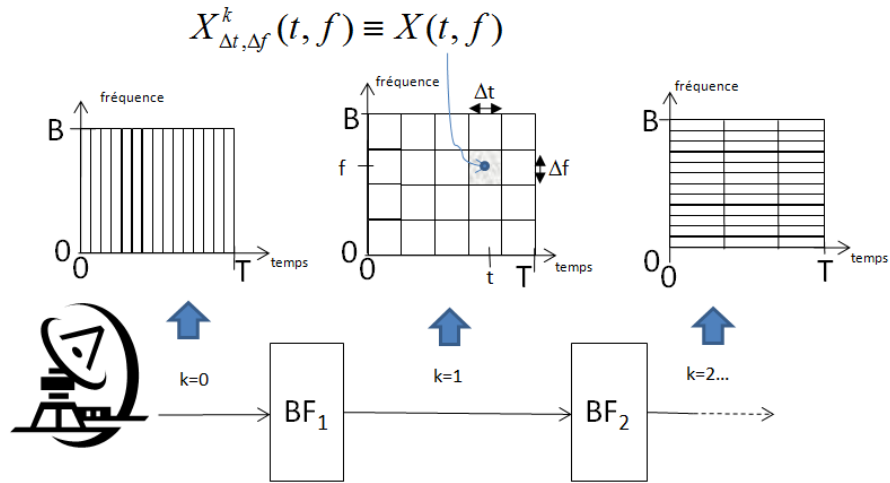


FIGURE 2.1 – Modélisation d'un récepteur de radioastronomie. Une succession de bancs de filtres (BF_k) fournit différents pavages temps-fréquence (T-F), $X_{\Delta t, \Delta f}^k(t, f)$ où k est un indice identifiant le banc de filtre, f et t sont respectivement les index fréquentiels et temporels qui positionnent le dit pavé T-F dans le plan T-F, Δf et Δt sont respectivement les résolutions fréquentielles et temporelles du pavage. Sauf besoin spécifique, le pavé T-F $X_{\Delta t, \Delta f}^k(t, f)$ sera simplement noté $X(t, f)$. Pour un signal initial de bande B , si nous considérons un point de mesure en début de chaîne, la résolution fréquentielle du plan T-F sera $\Delta f = B$ et la résolution temporelle $\Delta t = 1/B$. Par contre, en fin de chaîne, la résolution fréquentielle sera de $\Delta f = B/M$ et la résolution temporelle $\Delta t = M/B$ où M est le nombre de canaux fréquentiels que délivre la chaîne de mesure du radiotélescope.

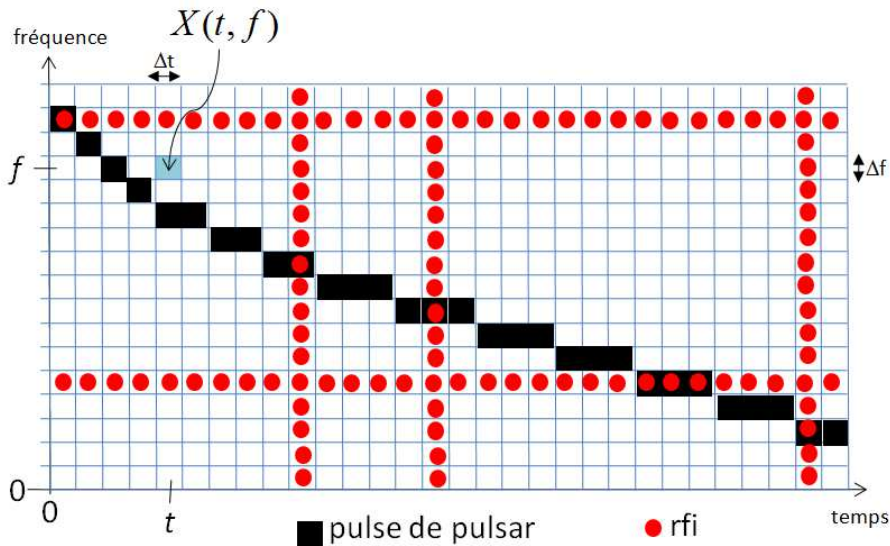


FIGURE 2.2 – Exemple de plan temps-fréquence (T-F). L'objectif des détecteurs de RFI est d'identifier les pavés T-F, $X(t, f)$, contenant une RFI (disque rouge). Pour un détecteur donné, le nombre de disques rouges détectés sur le nombre de disques rouges total s'appelle la probabilité de détection P_d et toute détection d'un pavé autre que celui contenant une RFI s'appelle une fausse alarme dont on peut également calculer la probabilité P_{fa} . Un très bon détecteur est celui qui maximise P_d tout en minimisant P_{fa} : l'idéal étant d'obtenir $P_d = 1$ et $P_{fa} = 0$.

si l'observation s'effectue avec un réseau d'antennes [Leshem and van der Veen, 1999] [Raza et al., 2002a] [Raza et al., 2002b] [Boonsta, 2004], [Hellbourg et al., 2012] ou à défaut l'utilisation d'une antenne de référence qui pointe vers la direction du brouilleur [Barnbaum and Bradley, 1998] [Poulsen et al., 2005]. Sinon, les méthodes d'estimation et soustraction peuvent être envisagées [Fridman, 2001] [Ellingson et al., 2001] [Bretteil, 2005]. Cependant, la plupart des instruments actuels étant encore à antenne unique, le filtrage spatial reste encore très anecdotique. Il ne sera d'ailleurs pas traité dans ce mémoire.

- Mélange nul : le signal radio cosmique et l'interférence ont au moins un support (temporel, fréquentiel, de polarisation) différencié. Théoriquement un filtrage adapté permet de séparer les deux sources. En pratique le filtrage fréquentiel est très utilisé. Alors qu'une séparation temporelle est peu envisageable puisque les sources cosmiques émettent souvent en continu. Le mélange nul peut se voir comme un cas particulier du mélange partiel.
- Mélange partiel : le signal radio cosmique et l'interférence partagent ponctuellement des mêmes zones temporelles et fréquentielles. C'est dans ce scénario que nous situons le cadre de cette thèse.

Si nous modélisons une chaîne de réception d'un radiotélescope comme une succession de bancs de filtres (cf. figure 2.1), nous obtenons en sortie de chaque banc de filtres un pavage du plan temps-fréquence (T-F). Considérons un pavé de ce plan T-F, $X(t, f)$, défini comme une information sur le signal à l'instant t dans le canal fréquentiel d'index f et de largeur Δf . La résolution temporelle Δt est telle que $\Delta f \cdot \Delta t = 1$.

Suivant les situations, le pavé T-F $X(t, f)$ peut représenter :

- une amplitude, $x(t, f)$ qui peut être complexe ou réel en fonction du contexte. Si on travaille à fréquence f fixée, on parlera alors plutôt d'échantillons, $x(t)$.
- la puissance instantanée du signal, $S(t, f) = \|X(t, f)\|^2$.
- ou tout autre combinaison, comme par exemple des corrélations croisées, des visibilitées....

L'objectif des différents détecteurs qui vont être cités dans la suite est de pouvoir établir la qualité *polluée* ou *non polluée* de ce pavé $X(t, f)$ (cf. figure 2.2). Une fois cette qualité établie par le ou les détecteurs, un traitement peut être envisagé, le plus simple étant de masquer/éliminer¹ ces pavés, des données exploitables par le radioastronome. Dans le cadre du projet Uniboard, des variantes à ce traitement basique sont proposées (cf. section 2.8).

Les performances des détecteurs de RFI qui vont être présentés dépendent évidemment de l'adéquation entre l'interférence et les propriétés statistiques (puissance, cyclostationnarité, non gaussianité) qu'ils évaluent. Mais ils dépendent aussi des caractéristiques temporelles et fréquentielles des RFI relativement aux caractéristiques temporelles et fréquentielles du pavage induit par les différents bancs de filtres de la chaîne de réception.

Ainsi, si nous considérons des détecteurs positionnés en bout de chaîne, comme cela se fait en pratique, les pavés T-F associés afficheront une très bonne résolution fréquentielle. Cela correspond à un cas optimum pour détecter des interférences à bande étroite. Par contre, ce pavage T-F sera tout à fait inadapté pour détecter des événements impulsifs (i.e. interférences à bande large). Pour ce dernier cas, il serait alors préférable de placer ces mêmes détecteurs en début de chaîne, là où la résolution temporelle est la meilleure.

Comme il n'est pas envisageable d'adapter la configuration de la chaîne de réception en fonction de l'interférence potentielle, l'option que nous proposons dans ce mémoire, c'est de positionner les détecteurs de RFI tout au long de la chaîne de réception du radiotélescope. Les simulations de la section 2.7 illustrent cette option suivant deux modalités d'implantation (implantation à durée physique constante ou implantation à nombre d'échantillons constant).

Avant d'aborder ces points, nous allons, dans la section suivante, présenter des techniques actuelles de détection fondées sur des critères de puissance.

2.3 Détecteurs d'interférences fondés sur le critère de puissance

2.3.1 Détecteur de radar

Le détecteur de radar (DetRadar) a été développé [Dumez-Viou, 2007] pour le NRT (section 1.2), dont les observations sont souvent brouillées par les systèmes Radars. Ces interférences sont de type large bande, intermittentes et émises avec une période de répétition de l'ordre de la milliseconde. L'algorithme est appliqué sur la mesure de la puissance instantanée, $S(t, f)$, à fréquence f fixée. On parle alors plutôt d'échantillons que de pavés T-F. L'algorithme est implantée en début de chaîne, c'est à dire lorsque la bande d'analyse, Δf est large (par exemple 14 MHz dans le cas de l'application sur le NRT).

1. on parle habituellement de blanking.

La première étape de l'algorithme est d'estimer une puissance moyenne de référence et cela de manière robuste (i.e. non affectée par la présence de RFI) et adaptative (i.e. capable de suivre les variations lentes de la puissance du signal utile). Cette estimation servira ensuite à fixer les seuils de détection. A f fixée, cet estimateur robuste, noté $\mu_N(t)$, est obtenu par une moyenne glissante des N dernières valeurs de $S(t, f)$ qui n'ont pas dépassé un seuil, noté Σ_{fort} . Une valeur typique est $\Sigma_{fort} = 4\mu_N(t)$.

A partir de ce schéma de base, deux règles sont proposées pour optimiser les performances de détection des impulsions radar fortes et faibles (cf. 2.3.a) :

- Le bloc "indicateur forte impulsion" : Pour des impulsions fortes, la détection est déclenchée si N_{fort} échantillons consécutifs dépassent le seuil fort Σ_{fort} . Par ce biais, on impose une certaine morphologie à la forme temporelle de la RFI, ce qui permet de réduire le taux de fausses alarmes.
- Le bloc "indicateur faible impulsion" : Pour des impulsions faibles, la détection est déclenchée si au moins N_{min} échantillons, dépassent un autre seuil, Σ_{faible} , dans une fenêtre de temps fixée à N_{faible} échantillons. $\Sigma_{faible} < \Sigma_{fort}$.

Les seuils Σ_{faible} et Σ_{fort} sont de la forme $\Sigma_x = C_x \cdot \mu_N(t)$ où les C_x sont 2 paramètres définis par l'utilisateur en fonction de niveau de fausse alarme désirée. Cet algorithme a été testé et implanté en temps réel pour les observations des raies d'hydrogène avec des résultats très satisfaisant comme le montre la figure 2.3.(b).

L'idée de fixer deux valeurs de seuils, une pour détecter les impulsions fortes et l'autre plus petite pour les impulsions faibles avec des fenêtres de temps à chaque fois différentes, est une idée généralisée par les travaux de [Offringa et al., 2010] qui sont présentés dans la section suivante.

2.3.2 Seuillage combinatoire : SumThreshold

Toujours dans le même objectif, qui est la détection de puissance robuste aux RFIs, [Offringa et al., 2010] propose de fixer plusieurs seuils adaptés à la longueur et à la puissance des RFIs qu'ils soient large bande ou bande étroite. Elle est donc appliquée dans les deux directions du plan T-F.

L'approche est la suivante : si deux pavés T-F voisins, par exemple $S(t, f)$ et $S(t \pm 1, f)$ ou $S(t, f)$ et $S(t, f \pm 1)$, ne dépassent pas individuellement un premier seuil Σ_1 , ils peuvent encore être marqués comme pollués si leur moyenne dépasse un deuxième seuil, Σ_2 , plus petit. Sinon, ils peuvent être moyennés à un troisième pavé T-F voisin et comparés à un troisième seuil, Σ_3 , encore plus petit et ainsi de suite jusqu'à arriver au dernier seuil appartenant à la suite des seuils fixés au début du traitement.

Pour éviter que les pavés T-F déjà détectés comme pollués dans les itérations précédentes imputent l'itération en cours, leur valeur est systématiquement remplacée par la valeur du seuil en cours.

Les seuils sont calculés avec la formule suivante :

$$\Sigma_k = \frac{\Sigma_1}{\rho^{\log_2(k)}} \quad k = 1, \dots, N_\Sigma \quad (2.1)$$

Avec ρ une valeur choisie empiriquement ($\rho > 1$, typiquement $\rho = 1.5$) et N_Σ , un entier donnant le nombre maximal de pavés T-F à combiner (typiquement, $N_\Sigma = 64$). Quant au premier seuil Σ_1 , il est calculé par simulation sur un jeu de données non polluées. Il doit permettre de ne pas dépasser la probabilité de fausse alarme souhaitée par l'utilisateur.

Pour des raisons d'efficacité, cette méthode est appliquée d'une manière itérative en incluant un lissage préalable du plan T-F. Ceci peut rendre son application aux observations pulsars, notamment pour des impulsions de forte intensité, peu adaptée.

Par contre elle affiche des performances très similaires à ce qui peut se faire manuellement. De ce fait elle est bien adaptée au traitement automatique de gros volume de données, la preuve en est que cette méthode est disponible en routine dans le flux de traitement, en bout de chaîne du radiotélescope LOFAR. Elle se généralise à tout type de plan T-F de type $\|X(t, f)\|$.

Le section suivante décrit une autre méthode de détection robuste aux interférences dédiée à la recherche de nouveaux pulsars.

2.3.3 Filtrage "zéro DM"

Cette méthode est basée sur le fait que les RFI à large bande (i.e. impulsionnelles) n'affichent pas une signature dispersée telle que la signature du pulsar (voir la section dispersion), d'où le nom filtrage "zéro-DM". Cet algorithme s'applique au flux de données avant l'étape de dédispersion dans le cadre de la recherche de nouveaux pulsars [Eatough et al., 2009].

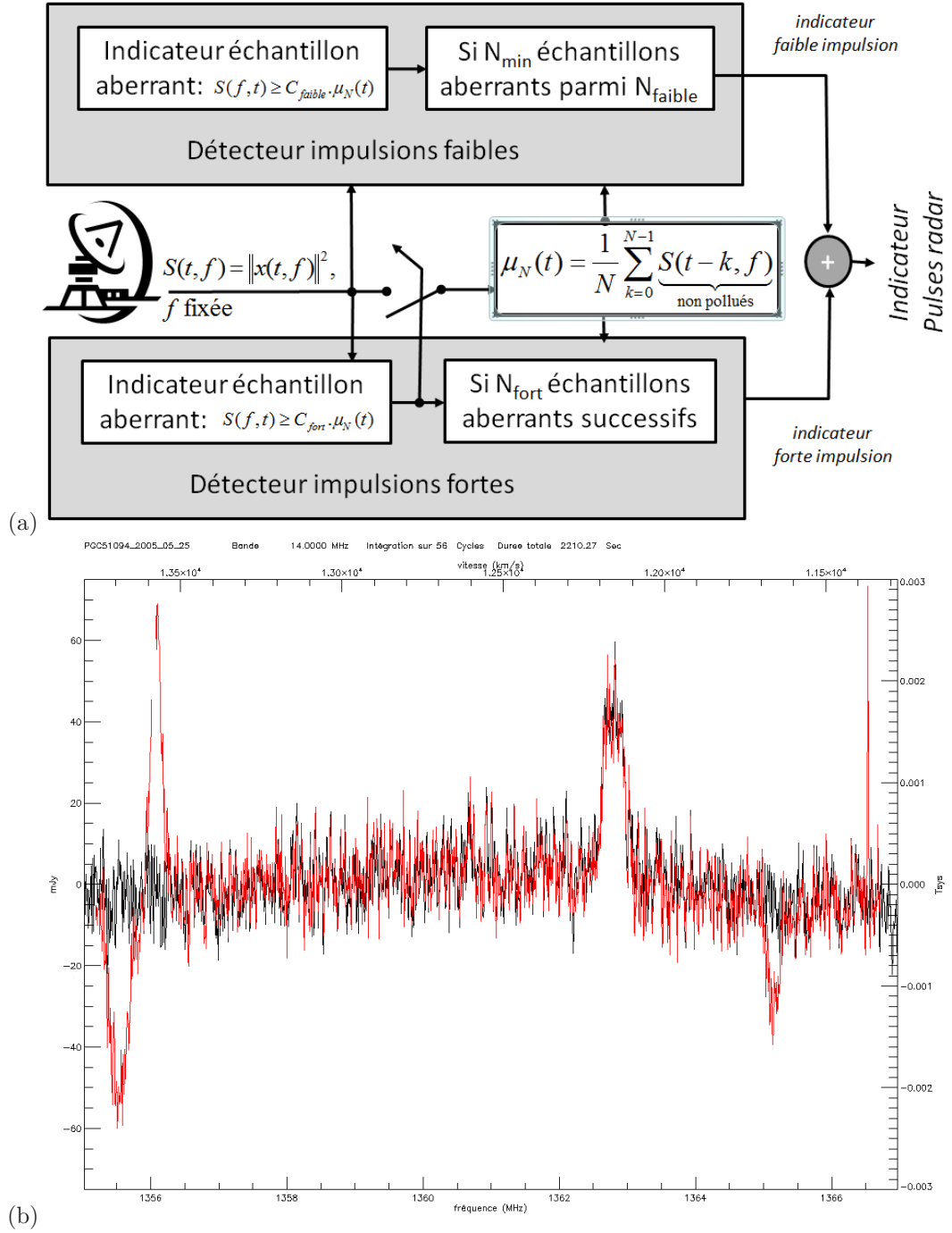


FIGURE 2.3 – Détecteur de puissance radar : (a) Schéma expliquant l'algorithme DetRadar. (b) Spectre de la raie 21 cm de l'hydrogène neutre (HI) de la galaxie PGC 51094 observée au NRT avec (noir) et sans (rouge) élimination radar. La figure a été tirée de [Dumez-Viou, 2007] avec son autorisation. Les paramètres de détection étaient les suivants : Bande d'analyse $\Delta f = 14$ Mhz, $N_{forte} = 3$, $N_{faible} = 30$, $N_{min} = 25$. Les probabilités de fausse alarme pour les indicateurs d'impulsions fortes et faibles étaient respectivement fixées à $2.1 \cdot 10^{-3}$ et $1.2 \cdot 10^{-3}$ via les paramètres respectifs $C_{forte} = 4$ et $C_{faible} = 0.8125$. La puissance de référence $\mu_N(t)$ était établie sur une fenêtre glissante de $N = 14000$ échantillons.

Le filtrage zéro-DM est mis en œuvre avant la dé-dispersion incohérente décrite dans la section 1.2.3, par un calcul simple de la moyenne, à un instant donné, t , sur les M canaux fréquentiels du plan T-F observé. Ainsi, un nouveau plan T-F, $S'(t, f)$, exempt de RFIs impulsionnels peut être construit :

$$S'(t, f) = S(t, f) - \frac{1}{M} \sum_{m=1}^M S(t, m) \text{ avec } f = 1, \dots, M \quad (2.2)$$

Il est clair que tout signal à bande large non dispersé entraînera une hausse de puissance simultanément dans tous les canaux et sera éliminé par ce processus.

Bien que ce filtrage induise in fine une diminution de l'intensité et une distorsion du profil du pulsar, les résultats présentés dans [Eatough et al., 2009] montrent que cette technique améliore d'une façon considérable la détection des impulsions individuelles de pulsar en minimisant les fausses détections du fait des RFIs impulsionnels.

En effet, une approche classique [Lorimer and Kramer, 2005] dans le cadre de recherche de nouveaux pulsars consiste à comparer chaque échantillon temporel dans un canal donné avec la moyenne plus deux ou trois fois l'écart type, de manière similaire à l'indicateur d'impulsions fortes du détecteur de radar (section 2.3.1). Si l'échantillon à comparer est plus grand que ce seuil, alors il est mis à zéro. Cette méthode nommée "DM zéro", à ne pas confondre avec le filtrage "zéro DM". D'après [Eatough et al., 2009], ceci peut provoquer la perte d'événements impulsionnels cosmiques peu dispersés et assez forts comme les impulsions géantes de pulsar. D'où l'intérêt du filtrage "zéro DM" qui repose sur l'hypothèse de la non dispersion des RFIs.

2.3.4 conclusions

Les principales méthodes fondées sur des critères de puissance viennent d'être présentées. Ces techniques sont intéressantes, notamment de par leur simplicité de mise en œuvre, si les impulsions de pulsar présentent des niveaux de puissance comparables à celui de référence des signaux naturels et si les RFIs sont nettement au dessus de ce niveau de référence. Sinon, des impulsions de pulsar plus importantes risquent d'être identifiées comme des RFIs. Réciproquement, des RFIs plus faibles risquent de ne pas être détectées.

Pour ce type de détecteur il est primordial de pouvoir établir la puissance de référence qui sert ensuite à définir les seuils de détection. Elle doit être obtenue de manière robuste, c'est à dire qu'elle doit être insensible aux contributions de RFIs. A ce propos, citons [Fridman, 2008] qui propose un tour d'horizon des estimateurs robuste de la puissance. Elle doit aussi être obtenue de manière adaptative, autrement dit pouvoir suivre les variations naturelles de la puissance, comme celles dues aux variations de température, aux mouvements de l'antenne, aux gabarits des filtres de la chaîne de réception. Toute imprécision sur cette puissance de référence va dégrader irrémédiablement les performances de détection [Sonnenschein and Fishman, 1992].

Le tableau 2.1 présente une comparaison récapitulative de ces méthodes.

Les sections suivantes présentent deux autres critères qui n'ont pas besoin de cette puissance de référence car ils sont fondés sur des caractéristiques statistiques intrinsèques aux RFIs et dont sont dépourvus les signaux naturels : la cyclostationnarité et la non-gaussianité.

Méthodes	Type de RFI	Robustesse	Adaptativité	Complexité	Où dans la chaîne de réception ?
DM zéro	impulsionnels	faible	aucune	très simple	en bout de chaîne de réception
DetRadar	dépend de la bande Δf	très bonne	très bonne, par moyenne glissante	Simple, approche temporelle (i.e. un détecteur par canal fréquentiel)	partout
SumThreshold	Impulsionnels et bande étroite	très bonne	très bonne, par lissage du plan T-F	lourde, Itérative sur tout le plan T-F	en bout de chaîne de réception
Filtrage "zéro DM"	Impulsionnels	sans objet	seulement dans le domaine temporel	Simple, orientée détection de nouveaux pulsars	en bout de chaîne de réception

TABLE 2.1 – Tableau comparatif des méthodes d'estimation au second ordre. A part le filtrage "zéro DM", ces techniques présentent toutes le risque de confondre une impulsion géante de pulsar avec une RFI impulsionnelle.

2.4 CycloDet : Détecteur de brouilleurs cyclostationnaires

2.4.1 Présentation du concept

La plupart des signaux de télécommunications présente une périodicité cachée due aux caractéristiques périodiques incluses dans la construction des signaux, comme par exemple la fréquence porteuse, la vitesse de transmission et le codage. Ces paramètres sont généralement cachés par le caractère aléatoire du message à transmettre. Toutefois, en utilisant une approche cyclostationnaire, qui fut introduite par [Gardner, 1987][Gardner, 1985], cette périodicité peut être récupérée. Vu que les signaux cosmiques peuvent être considérés comme stationnaires au second ordre, le critère de cyclostationnarité peut les séparer des signaux de télécommunications. Une vue exhaustive sur la théorie de la cyclostationnarité et ses applications peut être trouvée dans [Gardner et al., 2006] [Serpedin et al., 2005].

1. Principe de cyclostationnarité :

Considérons un canal fréquentiel donné, la fonction d'autocorrélation du signal $x(t)$ dans ce canal (on omet la fréquence pour plus de lisibilité) est définie par :

$$R_{x,x^*}(t, \tau) = \left\langle x\left(t + \frac{\tau}{2}\right)x^*\left(t - \frac{\tau}{2}\right) \right\rangle_{\infty}. \quad (2.3)$$

Où $\langle \cdot \rangle_N$ est la moyenne temporelle sur N échantillons. $\langle \cdot \rangle_{\infty}$ correspond donc à une moyenne asymptotique.

- Si $R_{x,x^*}(t, \tau)$ est indépendant de t le processus est dit *stationnaire*,
- Si $R_{x,x^*}(t + T_0, \tau) = R_{x,x^*}(t, \tau)$. On dit alors que $x(t)$ présente une périodicité de second ordre, et le processus est dit *cyclostationnaire* au second ordre. On définit la fréquence fondamentale correspondante et ses harmoniques comme *les fréquences cycliques*. Ces fréquences cycliques sont notées $\alpha = k/T_0$, avec $k \in \mathbb{Z}$.

Si $x(t)$ est un processus cyclostationnaire de période cyclique T_0 , alors $R_{x,x^*}(t, \tau)$ est T_0 -périodique et peut être développée en série de Fourier :

$$R_{x,x^*}(t, \tau) = \sum_{\alpha}^{+\infty} R_{x,x^*}^{\alpha}(\tau) e^{j2\pi\alpha t} \quad (2.4)$$

Avec $\alpha = \frac{k}{T_0}$ et $k \in \mathbb{Z}$, et où $R_{x,x^*}^{\alpha}(\tau)$ sont les coefficients de Fourier définis par :

$$R_{x,x^*}^{\alpha}(\tau) = \left\langle R_{x,x^*}(t, \tau) e^{-j2\pi\alpha t} \right\rangle_{T_0}. \quad (2.5)$$

$R_{x,x^*}^{\alpha}(\tau)$ est appelée la *fonction d'autocorrélation cyclique*. On peut aussi calculer la *fonction d'autocorrélation cyclique conjuguée* en remplaçant $x^*(t - \frac{\tau}{2})$ par $x(t - \frac{\tau}{2})$ dans l'équation 2.3.

2. La mise en oeuvre pratique :

En pratique, la fonction d'autocorrélation cyclique se calcule directement suivant cette formule :

$$R_{x,x^*}^{\alpha}(\tau) = \left\langle x\left(t + \frac{\tau}{2}\right)x^*\left(t - \frac{\tau}{2}\right)e^{-j2\pi\alpha t} \right\rangle_N. \quad (2.6)$$

Autrement dit, $R_{x,x^*}^{\alpha}(\tau)$ peut s'interpréter comme la transformée de Fourier sur N échantillons du produit de corrélation instantannée $x(t + \frac{\tau}{2})x^*(t - \frac{\tau}{2})$ et la distinction entre signal *pollué* et signal *non pollué* se fait simplement par le critère suivant :

- Si $R_{x,x^*}^{\alpha}(\tau)$ est nulle pour tout $\alpha \neq 0$, alors $s(t)$ est stationnaire et on considère que les N échantillons sont non pollués.
- Par contre, si $\exists \alpha \neq 0$, de sorte que $R_{x,x^*}^{\alpha}(\tau) \neq 0$, alors on a une signature due à la présence d'une RFI et on considère que les N échantillons sont pollués.

3. Le cas $\tau = 0$:

Remarquons que si on se limite au cas $\tau = 0$ alors le critère devient simplement :

$$R_{x,x^*}^{\alpha}(0) = \left\langle x(t)x^*(t)e^{-j2\pi\alpha t} \right\rangle_N = \left\langle S(t, f)e^{-j2\pi\alpha t} \right\rangle_N \quad (2.7)$$

Autrement dit, de simples transformées de Fourier sur chaque canal du plan T-F permettent de mettre en oeuvre le détecteur cyclostationnaire.

Remarquons également, que pour $\alpha = 0$, $R_{x,x^*}^{\alpha=0}(0)$ n'est que le calcul de la puissance moyenne dans un canal fréquentiel.

Avant de présenter les 2 détecteurs cyclostationnaires que nous proposons, nous allons examiner plus en détail le cas d'une RFI spécifique, la modulation de phase à 2 états (BPSK, Binary phase Shift Keying), car elle sera utilisée pour toutes les simulations par la suite.

2.4.2 Exemple de la modulation BPSK :

Nous avons considéré un signal BPSK (Binary Phase Shift Keying) comme étant le brouilleur qui nous intéresse pour les simulations car il présente des fréquences cycliques à la fois pour la forme directe et la forme conjuguée de la corrélation cyclique.

Notons $I(t)$ le signal issu d'une modulation BPSK dans un canal de fréquence donnée (pour simplifier les notations, nous ne notons plus le lien à un canal de fréquence) :

$$I(t) = \sum_k a_k h(t - kT_{sym} + t_0) e^{j(2\pi f_c t + \varphi)} \quad (2.8)$$

où les a_k sont des valeurs réelles prenant de manière aléatoire et indépendante l'amplitude $\pm\sigma_a$ (c'est le message binaire qui doit être transmis), $h(t)$ est la réponse impulsionnelle du filtre d'émission, $Br = 1/T_{sym}$ est le baud rate avec T_{sym} la durée d'un bit, t_0 est le retard, f_c et φ_0 sont respectivement la fréquence et la phase de la porteuse. La figure 2.4.a donne un exemple de $I(t)$.

Les fonctions d'autocorrélation cyclique et cyclique conjuguée pour le signal $I(t)$ sont données respectivement par les équations

$$R_{I,I^*}^\alpha(\tau) = \begin{cases} \frac{\sigma_a^2}{T_{sym}} e^{-j2\pi\alpha t_0} e^{j2\pi f_c \tau} r_{h,h^*}^{\frac{l}{T_{sym}}}(\tau) & \text{si } \alpha = \frac{l}{T_{sym}}, \\ 0 & \text{si } \alpha \neq \frac{l}{T_{sym}}. \end{cases} \quad (2.9)$$

$$R_{I,I}^\alpha(\tau) = \begin{cases} \frac{\sigma_a^2}{T_{sym}} e^{j2\varphi} e^{-j2\pi(\alpha - 2f_c)t_0} r_{h,h}^{\frac{l}{T_{sym}}}(\tau) & \text{si } \alpha = 2f_c + \frac{l}{T_{sym}}, \\ 0 & \text{si non.} \end{cases} \quad (2.10)$$

Avec $l \in \mathbb{Z}$ et

$$r_{h,h^{(*)}}^{\frac{l}{T_{sym}}}(\tau) = \int_{-\infty}^{+\infty} h(t + \frac{\tau}{2}) h^{(*)}(t - \frac{\tau}{2}) e^{-j2\pi \frac{l}{T_{sym}} t} dt \quad (2.11)$$

Le détail du calcul menant à ces deux équations est présenté dans [Gardner, 1991].

Ce qui va nous intéresser pour la suite ce sont les amplitudes de ces corrélations cycliques pour le cas $\tau = 0$. Les figures 2.4.c et 2.4.d tracent ces modules en fonction de la fréquence cyclique à partir d'une simulation comprenant du bruit gaussien et un signal BPSK. La figure 2.4.b donne un exemple sur une observation astronomique réelle. Les corrélations cycliques sont simplement calculées à partir de l'équation 2.7, c'est à dire par simple transformée de Fourier de $\|x(t)\|^2$ pour la corrélation cyclique et $x^2(t)$ pour sa version conjuguée, où $x(t)$ est l'amplitude mesurée dans un canal fréquentiel donné. Les positions des pics correspondent bien aux fréquences cycliques attendues, c'est à dire $\alpha = l/T_{sym}$ pour la fonction d'autocorrélation cyclique et $\alpha = 2f_c + l/T_{sym}$ pour la fonction d'autocorrélation cyclique conjuguée ($l \in \mathbb{Z}$). L'amplitude de ces pics est régie par le terme $r_{h,h^{(*)}}^{\frac{l}{T_{sym}}}(\tau)$ (Equ. 2.11) qui dépend de la réponse impulsionnelle du filtre d'émission, $h(t)$. Pour certaines modulations et pour certains filtres d'émission, le choix de $\tau = 0$ n'est pas optimum pour maximiser ces pics. Ce point est illustré en annexe B. Dans la suite du mémoire, nous considérons systématiquement le cas $\tau = 0$.

2.4.3 Les autres modulations :

Le tableau 2.2 donne les fréquences cycliques des corrélations cyclique et cyclique conjuguée pour différentes modulations. Pour le détail du calcul :

- la modulation d'amplitude (AM) est dans [Gardner, 1987] pages 400 et 401.
- BPSK [Gardner, 1991]
- La modulation de phase à K états et la modulation d'amplitude en quadrature (QAM) [Feliachi, 2010]

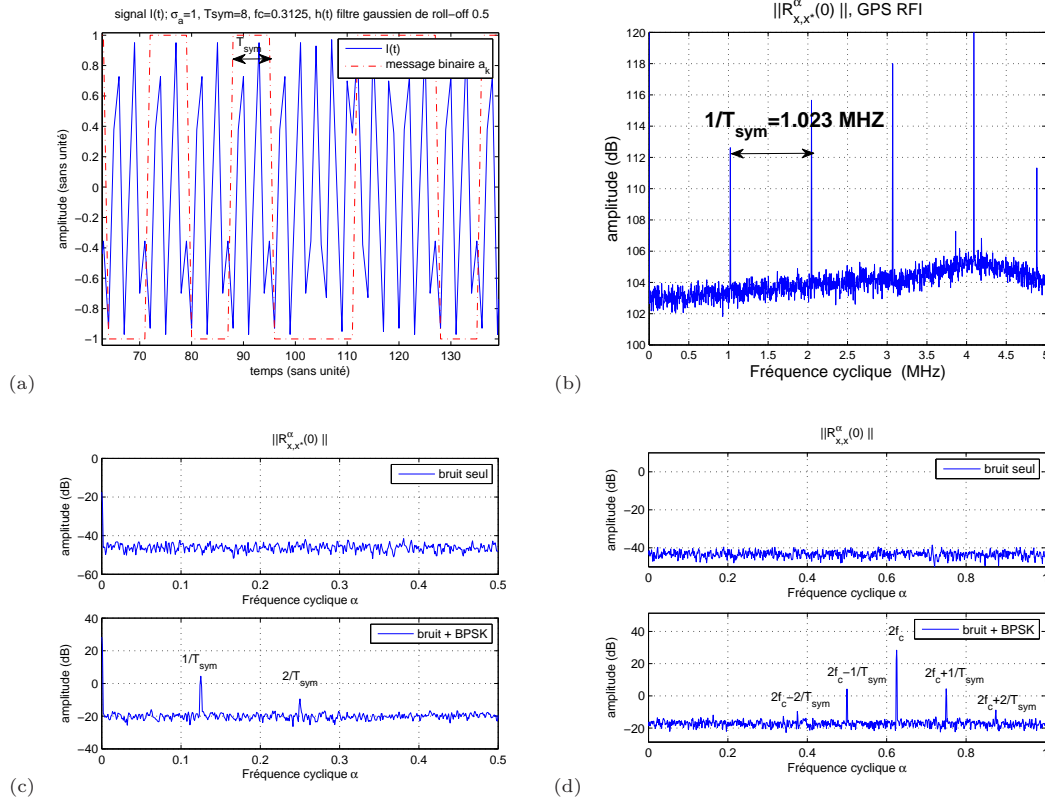


FIGURE 2.4 – La modulation BPSK et ses caractéristiques cyclostationnaires. (a) Représentation temporelle schématisée d'un signal BPSK. Les paramètres de ce signal BPSK sont ceux utilisés pour les figures c et d. (b) Cas cyclique, $\|R_{x,x^*}^\alpha(0)\|$, d'une observation au radiotélescope de Westerbork (Hollande) contenant un signal BPSK dû au satellite GPS. On retrouve la signature d'un des codes du GPS à 1023 MHz. (c) Cas cyclique $\|R_{x,x^*}^\alpha(0)\|$: (haut) bruit blanc gaussien seul, (bas) bruit blanc gaussien seul + BPSK (d) Cas cyclique conjugué $\|R_{x,x}^\alpha(0)\|$: (haut) bruit blanc gaussien seul, (bas) bruit blanc gaussien seul + BPSK. α est la fréquence cyclique normalisée, c'est à dire que $\Delta f = 1$ et $\Delta t = 1$.

- La modulation de fréquence à phase continue CPFSK [Agee et al., 1990] [Ferréol et al., 2004] [Gardner, 1987].
- La modulation à déplacement minimum gaussien (GMSK) [Agee et al., 1990] [Gardner, 1987] [Napolitano and Spooner, 2004].
- La répartition en fréquences orthogonales (OFDM) basée sur la OQAM (Offset Quadrature Amplitude Modulation) [Bölcskei, 2001] [Ciblat and Sempérin, 2003].

2.4.4 CycloDet et CycloDet 'aveugle' :

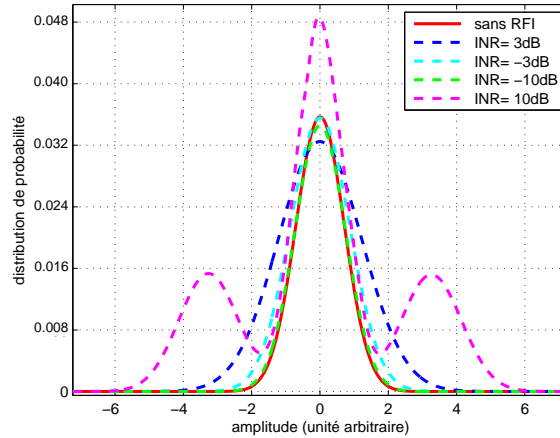
A partir des concepts présentés précédemment, nous proposons les critères de détection CycloDet, $D_{x^*}^\alpha$, et CycloDet conjugué, D_x^α , qui se définissent comme la corrélation cyclique ou cyclique conjuguée normalisée par rapport à la puissance du signal :

$$D_{x^*}^\alpha = \frac{|R_{x,x^*}^\alpha|}{R_{x,x^*}^0}. \quad (2.12)$$

La multitude de modulations utilisées de nos jours peut rendre compliquée la connaissance a priori des fréquences cycliques pour chaque signal de télécommunication. Mais surtout, les changements de fréquences et la succession des bancs de filtres dans la chaîne de réception rendent la détermination des paramètres de la modulation difficilement prédictible (mais pas impossible). Or ces paramètres sont nécessaires pour déterminer les fréquences cycliques à utiliser dans notre critère CycloDet.

Pour remédier à cette difficulté, nous proposons un critère aveugle ne nécessitant pas la connaissance a priori du type de modulation utilisée par les RFIs potentiels. L'idée est de calculer les critères définis par

Modulation	Fréquence cyclique	Fréquence cyclique conjuguée
AM	0	$2f_c$
BPSK	l/T_{sym}	$2f_c + l/T_{sym}$
MPSK et QAM	l/T_{sym}	0
CPFSK	l/T_{sym}	$2f_c + (2l+1)f_d$ $0 \leq l \leq \frac{m-2}{2}$
GMSK	l/T_{sym}	$2f_c \pm \frac{l}{2T_{sym}}$
OFDM et OQAM	$q/Q; q = 0, Q-1$	$\alpha_0 + q/Q; q = 0, Q-1$ $\alpha_0 = (\Delta f_c \bmod 1)$

TABLE 2.2 – Fréquences cycliques et fréquences cycliques conjuguées de différentes modulations ($l \in \mathbb{Z}$).FIGURE 2.5 – Distribution de probabilité d'un bruit blanc de puissance unité additionné à un signal BPSK (cf. Equ. 2.8) de différentes valeurs de puissance P_{BPSK} .

l'équation 2.12 pour un ensemble de fréquences cycliques via une Transformée de Fourier rapide (FFT). Il suffit ensuite de prendre le maximum de cette FFT pour disposer d'un critère qui fonctionne en aveugle.

Plus précisément, pour le détecteur cyclique aveugle :

$$D_{x^*} = \max_{\alpha=k/N} \left(\frac{|\langle x(t)x^*(t)e^{-j2\pi\alpha t} \rangle_N|}{\langle x(t)x^*(t) \rangle_N} \right) \quad (2.13)$$

avec $k = 0, \dots, N-1$.

Pour le détecteur cyclique conjugué aveugle, la formulation est similaire sauf qu'il faut remplacer $x^*(t)$ par $x(t)$ dans la partie supérieure de la fraction de l'équation précédente.

Pour ces deux détecteurs aveugles, le fait que les fréquences cycliques calculées ne correspondent pas forcément à la vraie valeur de la fréquence cyclique du RFI, a des répercussions sur les performances. Ce point a été étudié dans [Weber et al., 2007].

2.4.5 Conclusions

Dans cette partie, nous avons présenté le concept de cyclostationarité. Nous avons proposé deux détecteurs basés sur ce concept. Le premier suppose la connaissance a priori des caractéristiques de modulation de l'interférence et le second travaille en aveugle sans connaissance a priori. Avant d'entamer l'analyse des performances de ces détecteurs dans le cadre des observations de pulsar, nous allons présenter succinctement un autre critère de détection qui est lui aussi fondé sur une caractéristique intrinsèque aux RFIs : la non-gaussianité.

2.5 Estimateur spectral du Kurtosis

Comme l'illustre la figure 2.5, la non-gaussianité des RFIs est un critère potentiellement discriminant pour séparer signaux naturels et RFIs. L'estimateur du Kurtosis spectral (SK) a été initialement proposé par

[Nita et al., 2007], pour la détection et l'élimination des RFIs en temps réel sur un plan T-F en puissance (i.e. $S(t, f)$).

Considérant un plan T-F de M canaux fréquentiels et N instants temporels. Pour chaque canal fréquentiel $f = 0, \dots, M - 1$, on calcule la puissance instantanée $S(t, f)$, avec $t = 0, \dots, N - 1$. Pour chaque canal de fréquence f , le critère de détection défini par [Nita and Gary, 2010b], s'écrit :

$$SK(f) = \frac{N+1}{N-1} \left(\frac{NP_2(f)}{P_1^2(f)} - 1 \right) \quad (2.14)$$

avec

$$P_1(f) = \sum_{t=0}^{N-1} S(t, f) \quad P_2(f) = \sum_{t=0}^{N-1} S(t, f)^2 \quad (2.15)$$

Lorsque le plan T-F qui fournit les pavés $S(t, f)$ est issu d'un banc de filtres implanté sous la forme d'une simple transformée de Fourier, [Nita et al., 2007] a calculé les statistiques de ce critère. Il a ainsi montré qu'en l'absence de RFIs et en présence de signaux naturels réels ayant une statistique gaussienne, la moyenne de ce critère est égale à l'unité et sa variance est donnée par :

$$Var(SK(f)) \approx \begin{cases} \frac{4}{N} & f = 1, \dots, M/2 - 1 \\ \frac{24}{N} & f = 0, M/2 \end{cases} \quad (2.16)$$

D'autres travaux, qui ont pour but la généralisation de l'estimateur spectral de Kurtosis SK pour n'importe quelle loi de distribution, et notamment pour des signaux complexes sont disponibles dans [Nita and Gary, 2010a]. [Gary et al., 2010] a présenté le résultat de l'implantation de SK dans le cadre d'une expérimentation sur des signaux d'observations du Soleil en présence d'interférences. Ces résultats ont démontré que l'algorithme SK est particulièrement bien adapté pour identifier les RFIs intermittentes, mais qu'il aussi fonctionne pour certains types de RFI continues. Cet algorithme se caractérise en outre par sa facilité d'implémentation. Il est donc intéressant de le comparer à l'approche que nous proposons. C'est l'objet des sections suivantes.

2.6 Evaluation des techniques pour le chronométrage des pulsars

Dans cette section, l'objectif est, d'une part, d'évaluer la robustesse des deux classes de critères (cyclostationnaire et non-gaussien) en présence d'un pulse de pulsar, et d'autre part, d'étudier la sensibilité des dits détecteurs à la présence d'une RFI.

Ces évaluations se font avec la méthode de Monte Carlo (100 essais pour chaque scénario) sur la base d'un test d'hypothèses :

- pour le test de robustesse à un pulse pulsar :

$$\begin{cases} \text{hypothèse } H_0 : & x(t) = n(t) \\ \text{hypothèse } H_1 : & x(t) = n(t) + p(t) \end{cases} \quad (2.17)$$

- pour le test de sensibilité à une RFI de type modulation BPSK :

$$\begin{cases} \text{hypothèse } H_0 : & x(t) = n(t) + p(t) \\ \text{hypothèse } H_1 : & x(t) = n(t) + p(t) + I(t) \end{cases} \quad (2.18)$$

avec :

- $n(t)$ est un bruit complexe blanc gaussien centré de variance σ_n^2 . Il représente le bruit système de la chaîne de réception.
- $p(t)$ est l'impulsion pulsar. Cette impulsion est générée en utilisant le modèle fréquentiel donné par l'équation 1.17. La puissance de l'impulsion est σ_p^2 .
- $I(t)$ est un brouilleur BPSK défini par l'équation 2.8. Sa puissance est σ_I^2 .

On définit également :

- le rapport de la puissance de l'interférence sur la puissance du bruit en dB par :

$$RIB_{dB} = 10 \log_{10} \left(\frac{\sigma_I^2}{\sigma_n^2} \right), \quad (2.19)$$

- Le rapport de la puissance de l'impulsion pulsar sur la puissance du bruit par :

$$RPB_{dB} = 10 \log_{10} \left(\frac{\sigma_p^2}{\sigma_n^2} \right) \quad (2.20)$$

En outre, on adopte la terminologie suivante pour les différents détecteurs à tester :

- CycloDetBr : le détecteur cyclique à la fréquence cyclique $\alpha = 1/T_{sym} = Br$.
- CyclDetFc : le détecteur cyclique conjugué à la fréquence cyclique $\alpha = 2f_c$.
- CycloDetBrA : le détecteur cyclique aveugle.
- CycloDetFcA : le détecteur cyclique conjugué aveugle.
- SK : l'estimateur spectral du Kurtosis.

Tous ces détecteurs sont testés sur N échantillons d'un canal fréquentiel, f , étant fixé.

2.6.1 Robustesse des détecteurs à la présence d'un pulse de pulsar

1. Paramètres des simulations :

Dans cette section, on se place dans le scénario défini par l'équation 2.17. Paradoxalement, pour cette étude, un détecteur sera d'autant plus robuste et donc performant, qu'il ne sera pas capable de faire la différence entre les hypothèses H_0 et H_1 .

La figure 2.6 montre le plan T-F de $x(t)$ sous H_1 pour un $RPB = 10$ dB, à gauche une vue en deux dimensions et à droite une vue en trois dimensions. L'impulsion pulsar est bien visible. Pour toute les simulations de cette section, l'impulsion a un profil gaussien de largeur 64 échantillons. Elle est dispersée sur la totalité des $N = 16384$ échantillons générés.

2. Résultats :

Pour chaque détecteur, on trace le profil simplifié de la réponse du détecteur (PSRD) sous les hypothèses H_0 (bleue continue) et H_1 (rouge pointillée). La méthode d'obtention de ce profil est illustrée à la figure 2.7. La figure 2.8 représente les résultats obtenus avec les différents détecteurs, pour des RPB allant de -15 dB à 10 dB avec un pas de 5 dB.

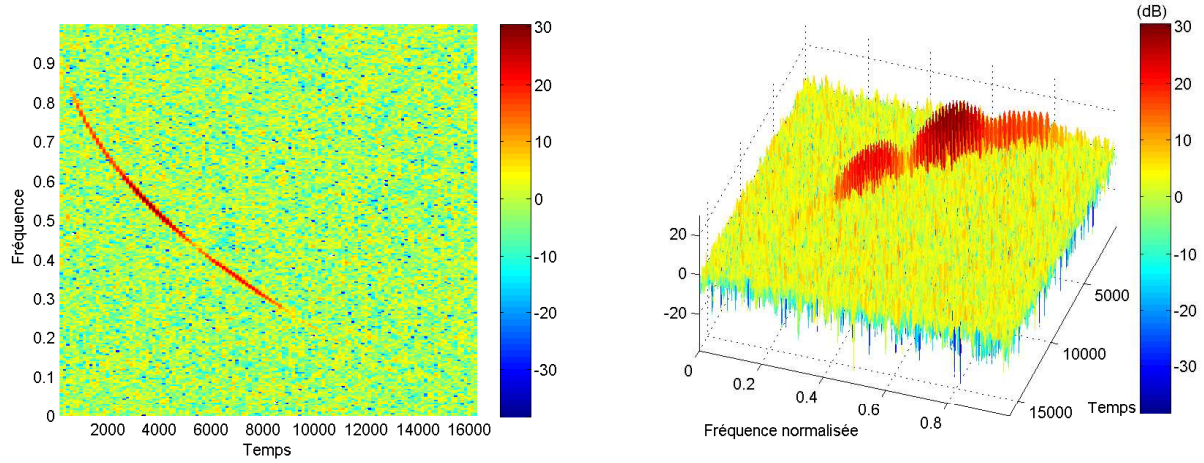


FIGURE 2.6 – Exemple de plan T-F de $x(t) = n(t) + p(t)$ avec $RPB = 10$ dB. Les axes des fréquences et du temps sont normalisés (i.e. $\Delta f = 1$ et $\Delta t = 1$). Les amplitudes sont en dB. L'impulsion a un profil gaussien de largeur 64 échantillons. L'impulsion est dispersée sur la totalité des $N = 16384$ échantillons générés.

- SK détecte systématiquement la présence de l'impulsion du pulsar à partir de -5 dB (voir la figure 2.8.(a) et son zoom 2.8.(b)).
- CycloDetFc et CycloDetBr ne sont pas sensibles à la présence du pulsar (figures 2.8.(c) 2.8.(e)).
- CycloDetFcA présente un risque de détection à partir de -5 dB (voir la figure 2.8.(d)). Ceci ne provient pas d'une quelconque nature cyclostationnaire de l'impulsion (il n'y a aucune raie dans le spectre cyclique) mais d'une déformation de la ligne de base du dit spectre cyclique du signal (cf. figure 2.9.(a)) qui induit une augmentation du détecteur sous H_1 .
- CycloDetBrA est plus robuste à cette déformation comme le montre la figure 2.10.(a). Pourtant, à $RPB = 10$ dB, sa variance augmente également fortement sous H_1 . Cela est dû à des détections sporadiques sous H_1 , comme le montre la figure 2.10.(c).

3. Conclusions :

Ces simulations montrent une meilleure robustesse de l'approche cyclostationnaire par rapport à l'approche non-gaussienne lorsqu'une impulsion de forte intensité est présente dans le jeu de données testé. Dans le cas des approches cyclostationnaires, la connaissance a priori des fréquences cycliques est également un plus par rapport aux détecteurs appliqués en aveugle (CycloDetBrA et CycloDetFcA).

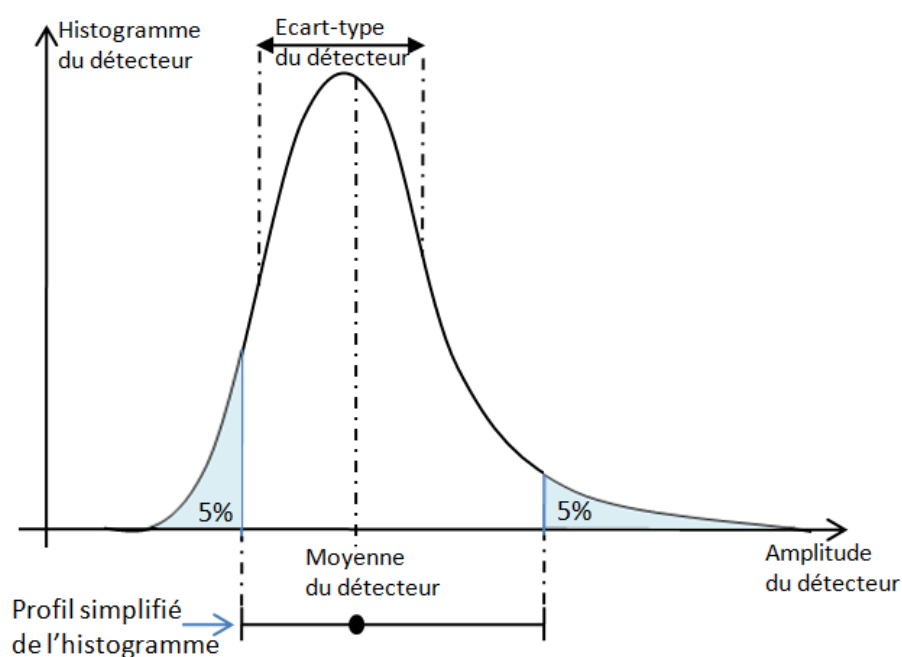


FIGURE 2.7 – Schéma de principe pour calculer le profil simplifié de la réponse du détecteur (PSRD). Considérons l'histogramme des valeurs fournies par un détecteur pour un scénario donné. Le profil simplifié de l'histogramme est donnée par la position de la moyenne et par la plage de valeurs obtenues hors des 5% valeurs extrêmes. Cette formulation permet de voir la symétrie et l'étalement de la distribution.

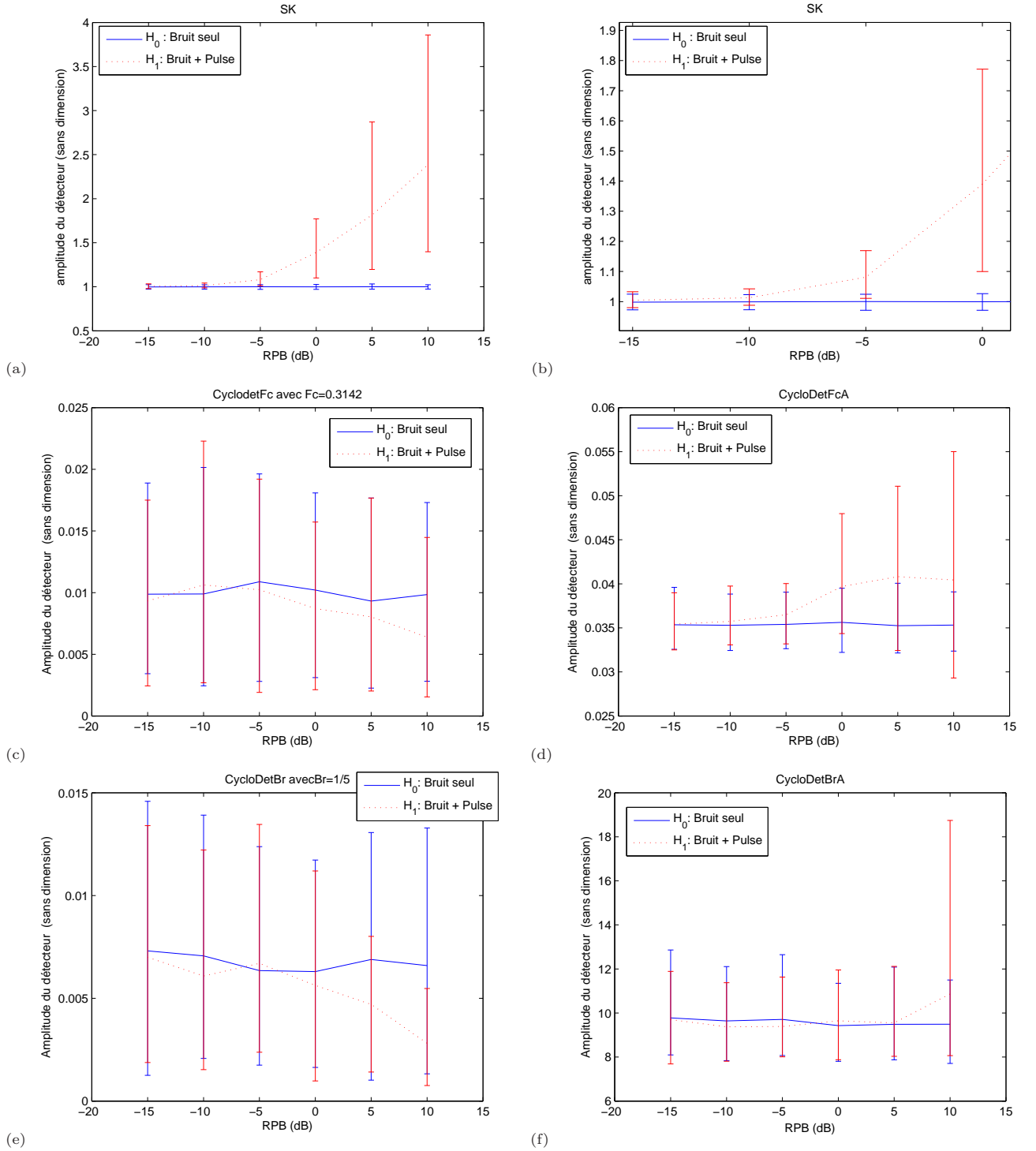


FIGURE 2.8 – Evaluation de la robustesse à la présence d'une impulsion pulsar. Les courbes représentent les profils simplifiés de l'histogramme des détecteurs sous H_0 et H_1 pour un RPB variable. L'impulsion a un profil gaussien de largeur 64 échantillons. Elle est dispersée sur la totalité des $N = 16384$ échantillons qui servent à la détection. (a) Résultat du détecteur SK. (b) Zoom de (a). (c) Résultat de CycloDetFc. (d) Résultat de CycloDetFcA. (e) CycloDetBr. (f) CycloDetBrA.

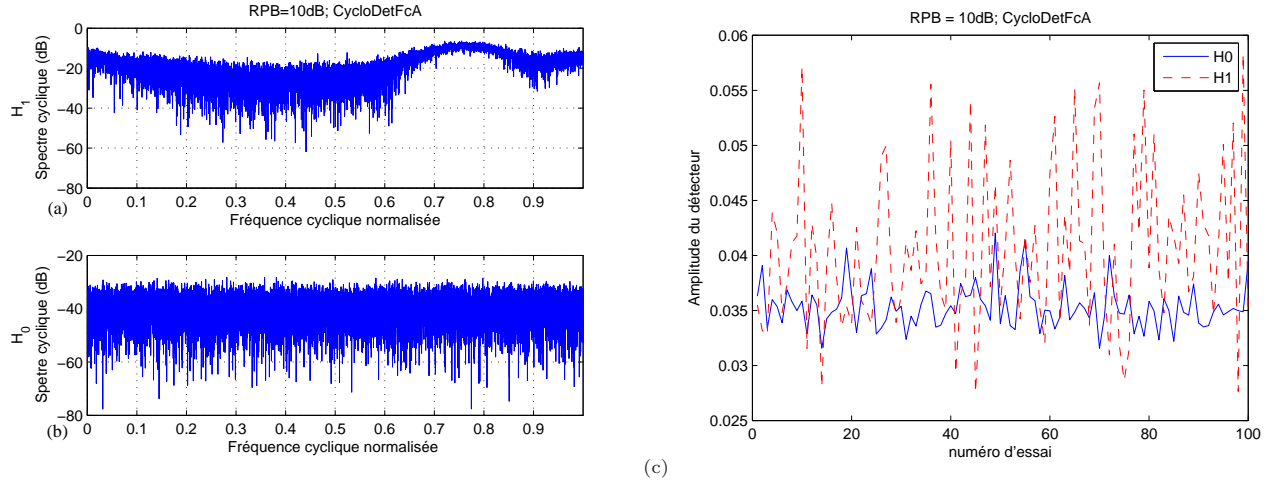


FIGURE 2.9 – (a) spectre cyclique, $\|R_{x,x}^\alpha(0)\|$, sous H_1 . (b) spectre cyclique sous H_0 . (c) 100 valeurs de CycloDetFcA sous H_0 et H_1 .

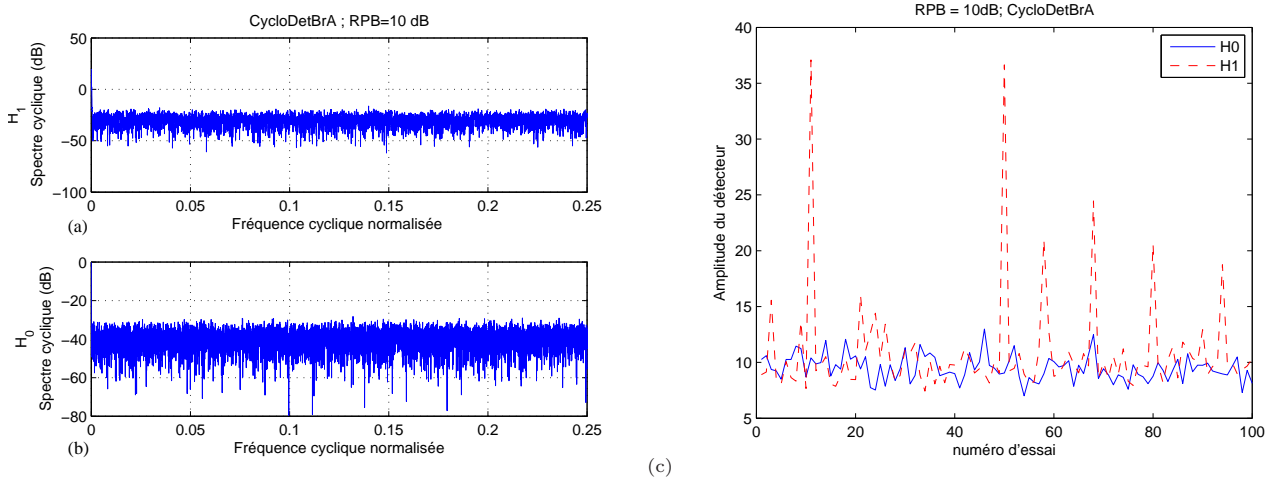


FIGURE 2.10 – (a) spectre cyclique, $\|R_{x,x^*}^\alpha(0)\|$, sous H_1 . (b) spectre cyclique sous H_0 . (c) 100 Valeurs de CycloDetBrA sous H_0 et H_1 .

2.6.2 Sensibilité des détecteurs en présence de RFI de type BPSK

Dans cette section, nous nous plaçons dans le scénario défini par l'équation 2.18. Comme précédemment, l'impulsion a un profil gaussien de largeur 64 échantillons, elle est dispersée sur la totalité des N échantillons générés et elle est mélangée à un bruit complexe blanc gaussien de telle sorte que le RPB soit fixé à 0 dB. Les RFIs sont de type BPSK comme défini par l'équation 2.8. En jouant sur le baud rate Br , nous considérerons 2 types de RFI, une à bande étroite et une à bande large, donc plus étalée. La figure 2.11 montre un exemple de plans T-F du signaux obtenus sous H_1 pour un $RIB = -13$ dB.

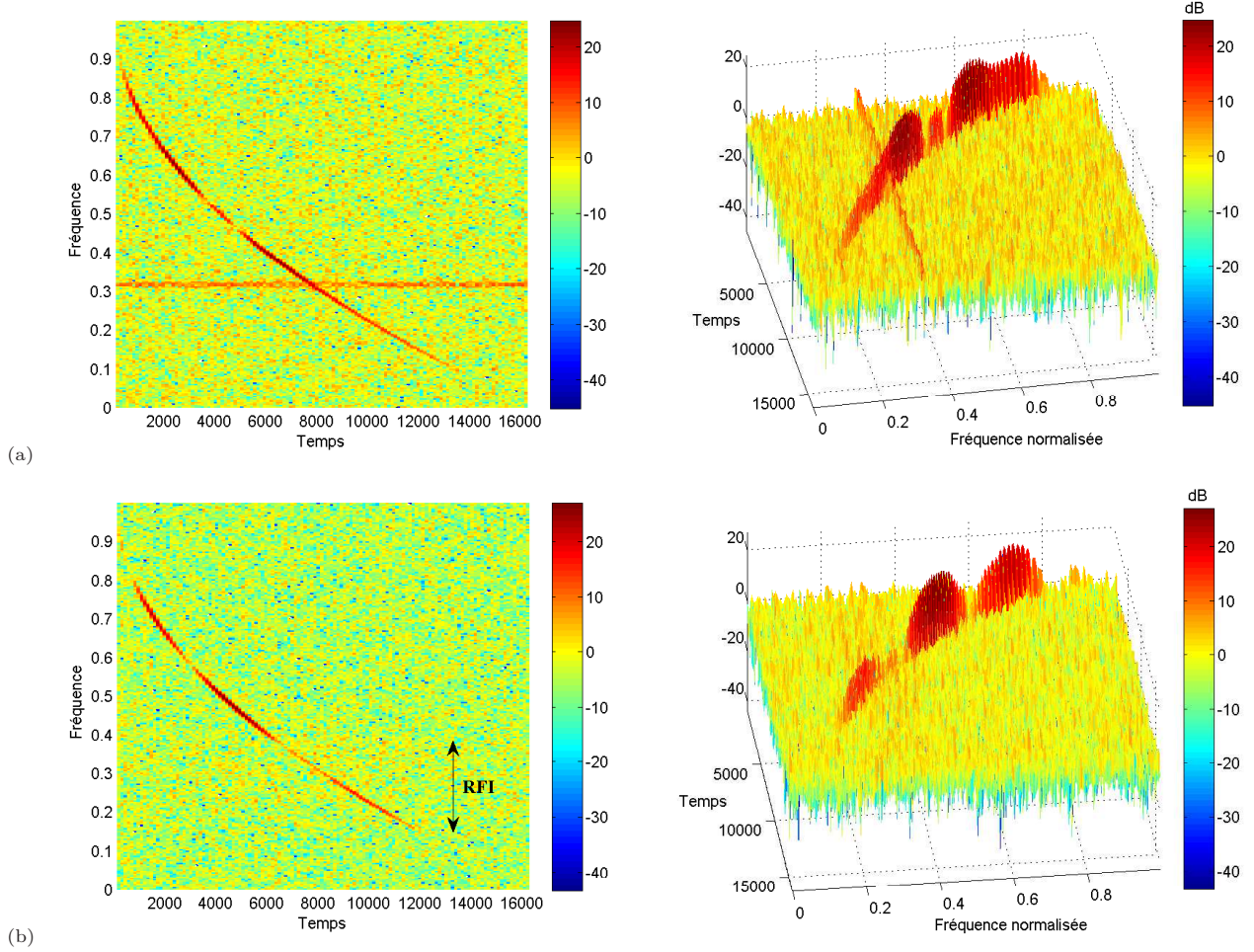


FIGURE 2.11 – Exemples de RFIs de type BPSK mélangés à une impulsion pulsar et un bruit blanc gaussien. Les paramètres communs à cette simulation sont $N = 16384$, $RPB = 0dB$, $RIB = -13dB$ et $F_c = 0.3142$. Les figures représentent le plan T-F $S(f, t)$ du signal en 2D (figures de gauche) et en 3D (figures de droite) (a) Cas du BPSK bande étroite, $Br = 1/100$ (i.e. un bit dure 100 échantillons) (b) Cas du BPSK bande large, $Br = 1/5$ (i.e. un bit dure 5 échantillons). Ici l'étalement du spectre de la RFI fait qu'elle est à peine visible en bas du plan T-F sous la forme d'une bande entre 0.2142 et 0.4142.

1. **Le critère de Fisher :** Pour évaluer les performances de ces détecteurs, les critères du second degré, s'exprimant en terme de moyenne et de variance des valeurs fournies par les détecteurs, sont souvent employés [Gardner, 1980]. Un critère du second ordre possible est le critère de Fisher [Fisher, 1936] [Fisher, 1938]. Il est défini par la formule :

$$F = \frac{(E_{H_1}[D] - E_{H_0}[D])^2}{Std_{H_1}[D]^2 + Std_{H_0}[D]^2} \quad (2.21)$$

où D est le détecteur sous test, $E[.]$ et $Std[.]$ sont respectivement l'espérance mathématique et l'écart-type sous H_0 ou H_1 . Concrètement, à partir de l'histogramme, obtenu par Monte Carlo, de la réponse du

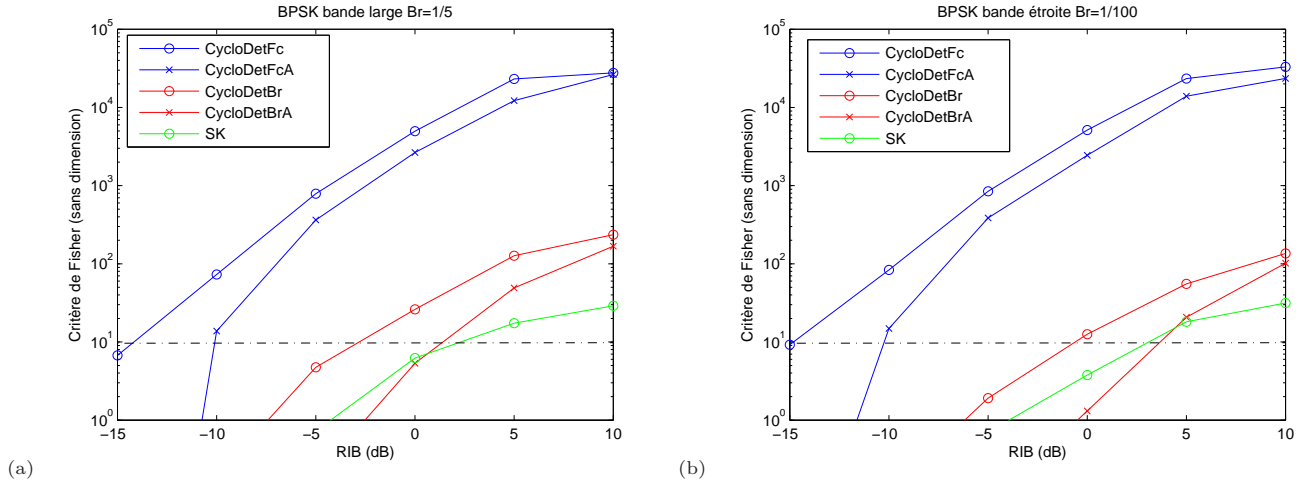


FIGURE 2.12 – Test de Fisher des détecteurs en fonction de RIB . Les paramètres communs de la simulation sont $N = 16384$, $RPB = 0dB$, $Fc = 0.3142$. Le Monte-Carlo s’est fait sur 100 essais. La ligne horizontale donne la limite pour obtenir une détection parfaite du RFI. (a) Cas du BPSK bande large avec $Br = 1/5$. (b) Cas du BPSK bande étroite avec $Br = 1/100$.

détecteur, on calcule la moyenne et l’écart-type pour chaque hypothèse (cf. figure 2.7) et on applique la formule ci-dessus.

En d’autres termes, le critère de Fisher permet de mesurer le contraste entre les deux hypothèses H_0 et H_1 . Ainsi, plus F est élevé, plus H_0 et H_1 sont séparées. Des valeurs de Fisher au delà de 10 garantissent une détection parfaite, c’est à dire qu’on peut trouver un seuil qui sépare parfaitement H_0 et H_1 . Dans ce cas, on obtient une probabilité de détection $P_d \approx 1$ et une probabilité de fausse alarme $P_{fa} \approx 0$. A contrario, des valeurs de F inférieures à 1 démontrent un détecteur totalement inefficace (i.e. pas de séparation entre H_0 et H_1). Entre les deux, on peut ajuster le compromis entre P_d et P_{fa} en jouant sur la valeur du seuil de décision.

2. **Résultats :** Nous avons analysé la sensibilité des différents détecteurs sous deux angles. D’une part nous avons considéré la sensibilité à N fixé ($N = 16384$) pour différentes valeurs du RIB de $-15 dB$ à $10 dB$ avec un pas de $5 dB$. Cela correspond à la figure 2.12. Et d’autre part, nous avons considéré la sensibilité à RIB fixé ($RIB = 0dB$) pour différentes valeurs de N de 256 à 16384 échantillons. Cela correspond à la figure 2.13.

- les détecteurs cyclostationnaires conjugués (CycloDetFc et CycloDetFcA) sont nettement plus performants que les autres. Ils peuvent détecter des RFIs plus faibles ou avec moins d’échantillons que les autres détecteurs testés. Cela vient du fait que pour le BPSK, l’amplitude des fonctions d’autocorrélations cycliques (cf. Equ.2.9) et cycliques conjuguées (cf. Equ.2.10) dépend du terme $r_{h,h(*)}^{\frac{l}{T_{sym}}}(0)$ (cf. Equ.2.11) qui lui est maximal pour $l = 0$. Or le cas $l = 0$ n’est possible que pour la fréquence cyclique conjuguée $\alpha = 2Fc$ (cf. annexe B).
- Pour les détecteurs fondés sur la cyclostationnarité, l’option aveugle est toujours moins performante que l’option avec connaissance a priori des fréquences cycliques.
- Les détecteurs cyclostationnaires fondés sur la fréquence cyclique $\alpha = Br$ sont sensibles à l’aspect large bande.

3. Conclusions :

Ces simulations montrent une meilleure sensibilité aux RFIs de l’approche cyclostationnaire par rapport à l’approche non-gaussienne.

Dans le cas des approches cyclostationnaires, la connaissance a priori des fréquences cycliques est également un plus par rapport aux détecteurs appliqués en aveugle (CycloDetBrA et CycloDetFcA).

L’aspect bande étroite ou bande large du RFI à l’intérieur du canal fréquentiel testé n’a qu’un impact limité sur la sensibilité. A la section 2.7, nous regarderons ces aspects lorsque ces caractéristiques fréquentielles du RFI s’étalent sur plusieurs canaux du bancs de filtres.

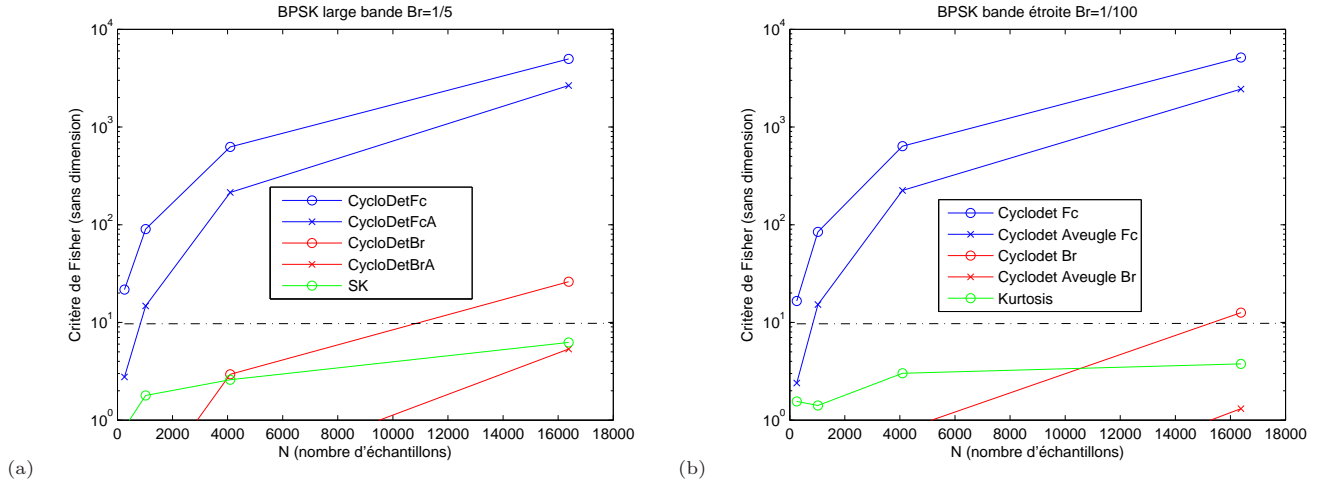


FIGURE 2.13 – Test de Fisher des détecteurs en fonction de N . Les paramètres communs de la simulation sont $RPB = 0dB$, $RIB = 0dB$, $Fc = 0.3142$. Le Monte-Carlo s'est fait sur 100 essais. La ligne horizontale donne la limite pour obtenir une détection parfaite du RFI. (a) Cas du BPSK bande large avec $Br = 1/5$. (b) Cas du BPSK bande étroite avec $Br = 1/100$.

2.6.3 Sensibilité des détecteurs après dédispersion

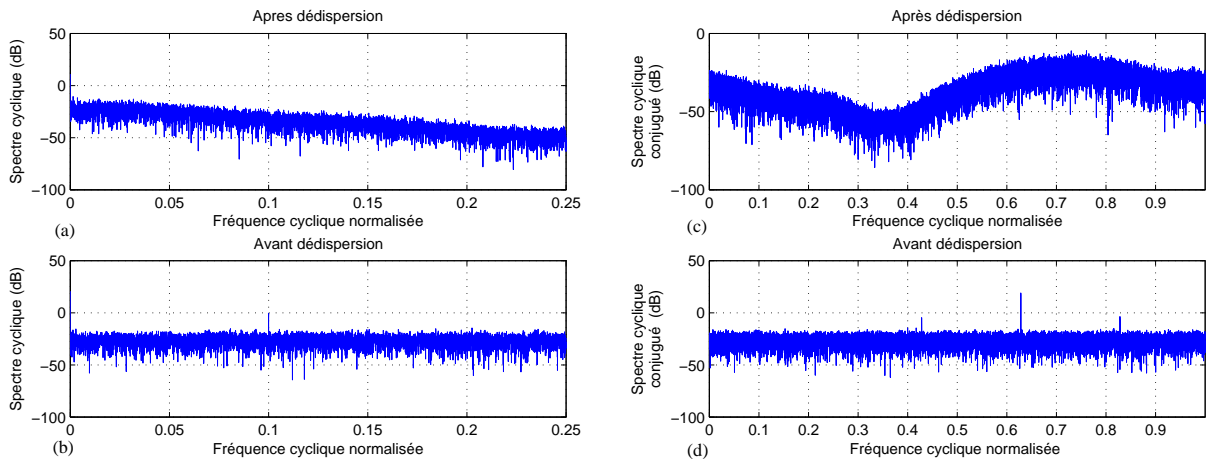


FIGURE 2.14 – Spectres Cycliques et spectres cycliques conjugués avant et après dédispersion du signal. Les figures du haut correspondent aux résultats après dédispersion pour le spectre cyclique (a) et le spectre cyclique conjugué (c). Les figures du bas correspondent aux résultats avant dédispersion pour le spectre cyclique (b) et le spectre cyclique conjugué (d).

Dans cette partie, nous avons voulu étudier si la mise en oeuvre des détecteurs fondés sur la cyclostationnarité ou la non-gaussianité était possible après le processus de dédispersion cohérent utilisé dans le chronométrage de pulsar.

Les simulations réalisées montrent que les 5 détecteurs étudiés perdent leurs capacités de détection après la dédispersion. Ceci est expliqué par deux points. D'une part, l'impulsion qui était avant dispersée devient non dispersée et le RPB augmente considérablement, ce qui déforme le spectre cyclique comme sur la figure 2.6.1. D'autre part, les RFI perdent leurs caractéristiques statistiques à cause de la dédispersion. La figure 2.14.(a) montre le spectre cyclique obtenu après dédispersion pour un BPSK à $Br = 1/5$ et $RIB = 10dB$. Il n'y a plus les raies aux fréquences cycliques attendues. La figure 2.14.(b) redonne le spectre cyclique du même signal avant dédispersion pour comparaison. La même constatation est faite pour le spectre cyclique conjugué après dédispersion (voir les figures 2.14.(c) et (d)).

Pour le bon fonctionnement des détecteurs, la possibilité de les implanter après l'étape de dédispersion n'est

pas envisageable.

2.7 L'influence du filtrage sur les performances de détecteurs

L'objectif de cette section est double :

- C'est une évidence de dire qu'un détecteur quel qu'il soit sera d'autant plus performant que le RIB de la RFI sera élevé. Une autre évidence est que nous n'avons aucun contrôle sur la puissance intrinsèque de la RFI qui est captée par un radiotélescope. Donc la seule manière d'augmenter le RIB c'est de diminuer le bruit présent avec la RFI.
Autrement dit, dans une chaîne de réception de radioastronomie, il faudrait pouvoir adapter la résolution fréquentielle, Δf , des canaux à la largeur spectrale des RFIs afin d'atteindre des RIB optimaux. Dans le cadre du projet européen Uniboard qui sera décrit à la section 2.8, il est prévu de mettre en place une chaîne de réception telle que celle schématisée à la figure 2.1. La succession des bancs de filtres (BF) permettrait alors d'accéder à des résolutions en temps et en fréquence variables. Les détecteurs en sortie de ces bancs de filtres auraient alors accès à différentes valeurs de RIB de la même RFI. Un objectif de cette partie est d'évaluer le comportement des détecteurs que nous avons étudiés dans une telle configuration.
- Le second objectif concerne l'impact du filtrage successifs sur les caractéristiques "cyclostationnarité" et "non gaussianité" des RFIs. En effet, dans les simulations précédentes, nous avons considéré que la RFI, qu'elle soit bande étroite ou large bande, se trouvait pleinement à l'intérieur d'un des canaux en sortie du banc de filtres. La question que nous posons maintenant est que deviennent les performances des détecteurs lorsque les caractéristiques fréquentielles de la RFI deviennent plus larges que la résolution fréquentielle Δf des bancs de filtres ?

Le protocole proposé pour cette simulation repose sur les points suivants :

- On considère 8 bancs de filtres successifs, chacun fournissant une résolution fréquentielle deux fois meilleure que le banc de filtres précédent. On a donc accès à 9 résolutions fréquentielles données par la formule suivante :

$$\Delta f_k = \frac{B}{2^k} \quad k = 0, \dots, 8 \quad (2.22)$$

avec B la bande d'entrée de la chaîne de réception. En entrée de ce système, il y a un seul canal de largeur B et en sortie, il y a $N_c = 256$ canaux de largeur $\frac{B}{256}$. La résolution temporelle suit évidemment la loi inverse :

$$\Delta t_k = \frac{2^k}{B} \quad k = 0, \dots, 8 \quad (2.23)$$

- dans l'optique d'une implantation matérielle à terme pour le projet Uniboard, deux options d'implantation vont être testées : implantation à nombre d'échantillons N constant et implantation à durée physique T constante où $T = N \cdot \Delta t_k$.
- L'implantation à *nombre d'échantillons N constant* suppose que quelque soit la position $k = 0, \dots, 8$ du détecteur, la détection se fait sur un nombre d'échantillons N fixe. Cependant, cela signifie que la durée d'observation en temps physique T est croissante en 2^k . L'intérêt d'une telle configuration est de conserver la même sensibilité de détection tout au long de la chaîne. N sera pris égal à 256 échantillons dans toutes les simulations présentées.
- A contrario, l'implantation à *durée physique T constante* va imposer une variation du nombre d'échantillons utilisés par le détecteur, sous la forme $N = \frac{T \cdot B}{2^k}$. L'intérêt d'une telle configuration est de conserver le même bloc T-F tout au long de la chaîne. La synchronisation des différents détecteurs en sera alors nettement facilitée, si l'implantation matérielle de cette option est retenue dans le projet Uniboard. T sera pris égale à $\frac{65536}{B}$ dans toutes les simulations présentées. Autrement dit, en fréquence normalisée (i.e. $B = 1$), les détecteurs auront accès à $N = 65536$ échantillons avant le premier banc de filtres ($k = 0$) et à $N = 256$ échantillons après le dernier banc de filtres ($k = 8$).
- Les résultats des simulations seront condensés sous la forme d'une carte des différents pavés T-F à chaque niveau de filtrage. Chaque pavé sera affecté par la valeur du critère de Fisher calculé sur un Monté-Carlo de 100 essais. Les valeurs du critère de Fisher inférieures à 1 sont bloquées à 1. La figure 2.15 récapitule le sens à donner à ces cartes.
- les bancs de filtres sont implantés par des FFT pondérées par un fenêtrage de Hanning. Il n'y a pas de recouvrement entre les FFTs donc cela correspond à une décimation maximale en sortie du banc de filtres.
- Du fait de la décimation et du repliement qu'elle induit, il est compliqué de prédire les fréquences cycliques attendues en sortie de chaque banc de filtres. CycloDetFc et CycloDetBr ne peuvent donc pas être appliqués simplement. Les détecteurs comparés dans cette partie ne seront donc que SK, CycloDetFcA et CycloDetBrA.

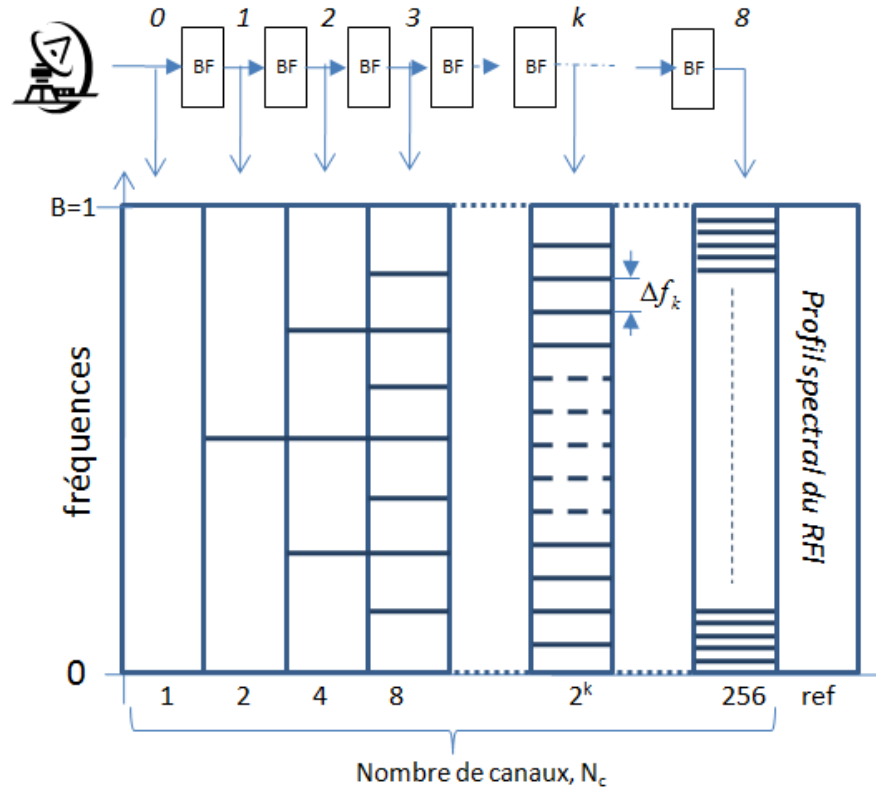


FIGURE 2.15 – Schéma du protocole de test. Chaque pavé sera affecté par la valeur du critère de Fisher du scénario considéré. La sortie du banc de filtre (BF) numéro k contiendra $N_c = 2^k$ canaux de résolution fréquentielle, $\Delta f_k = \frac{1}{2^k}$ avec $k = 0, \dots, 8$. La dernière colonne indexée "ref" représente le profil spectral de la RFI seule. Elle permet de vérifier visuellement l'adéquation de ce profil avec les résultats de détection dans les différents pavés.

1. **Cas du pulsar seul :** Le scénario considéré est celui de l'équation 2.17. L'impulsion pulsar a un profil gaussien de largeur 64 échantillons. Elle est dispersée sur la totalité des échantillons générés en entrée du système ($k = 0$) avec un $RPB = 10dB$. Les résultats sont donnés à la figure 2.16.

On remarque que le détecteur SK reste peu robuste à la présence d'une impulsion de pulsar. La détection est systématique quelque soit le nombre de canaux N_c (la valeur du test de Fisher est supérieur à 1) et s'accroît à partir de $N_c = 16$. CycloDetFcA ne détecte pas le pulse pour $N = 1, 2, 4$. Il l'a détecté à partir de $N_c = 8$. Quant à CycloDetBrA, la détection commence à partir de $N_c = 32$.

Ceci est expliqué par le fait que pour $N = 1$ l'impulsion est bien dispersée sur l'ensemble des points. Alors que pour $N = 32$ par exemple, l'impulsion occupe quelques échantillons avec une grande puissance ; SK est très sensible à cela, les détecteurs cyclostationnaire le deviennent aussi car cela déforme le spectre cyclique.

2. **sensibilité aux RFIs, cas d'un brouilleur très faible :**

Le scénario considéré est celui de l'équation 2.18.

Les paramètres communs des différentes simulations sont $Fc = 0.31416$, $RIB = -20 dB$, $RPB = -20 dB$. Seul le détecteur CycloDetFcA est utilisé car nous avons montré dans les sections précédentes que c'était le plus sensible. Quelque soit l'approche (T constant ou N constant), le brouilleur n'est détecté (voir la figure 2.17) que lorsqu'il y a adéquation entre les caractéristiques spectrales de la RFI et la résolution spectrale du banc de filtres. Cet exemple illustre l'intérêt d'avoir une approche multi-résolution pour la détection automatique des RFIs.

3. **sensibilité aux RFIs, cas d'un brouilleur fort :**

Le scénario considéré est celui de l'équation 2.18.

Les paramètres communs des différentes simulations sont $Fc = 0.31416$, $RIB = 10 dB$, $RPB = -10 dB$. 3 largeurs spectrales différentes pour la RFI vont être simulées : une RFI de type BPSK en bande étroite ($Br = 1/100$), une RFI de type BPSK en bande large ($Br = 1/5$) et une RFI de type BPSK avec une largeur de bande intermédiaire ($Br = 1/37$).

Les brouilleurs étant fort ($RIB = 10 dB$), l'effet de l'adéquation entre les caractéristiques spectrales de la RFI et la résolution spectrale du banc de filtres ne sera pas visible ici. L'objectif de cette simulation est donc essentiellement d'étudier comment les propriétés de cyclostationnarité et de non-gaussianité sont conservées (ou non) au fur et à mesure des filtrages.

Les résultats obtenus sont donnés à la figure 2.18 pour $Br = 1/100$, à la figure 2.19 pour $Br = 1/37$ et à la figure 2.20 pour $Br = 1/5$.

Le brouilleur bande étroite ($Br = 1/100$) étant pratiquement tout le temps dans une configuration où son lobe principal est à l'intérieur d'un canal fréquentiel, ses caractéristiques cyclostationnaires et de non-gaussianité sont conservées.

Lorsque les caractéristiques spectrales de la RFI s'élargissent, $Br = 1/37$ puis $Br = 1/5$, les filtres ayant une résolution spectrale trop fine mettent en défaut les détecteurs. Sachant que $2Br$ représente approximativement la largeur du lobe principal d'une modulation BPSK, on peut prédire que les performances des détecteurs commenceront à diminuer à partir de $k = 4$ pour $Br = 1/37$ et $k = 2$ pour $Br = 1/5$. C'est ce qu'indiquent également les simulations.

Il est intéressant de noter que lorsque qu'on isole le signal autour des maximum des lobes secondaires des RFIs, des caractéristiques cyclostationnaires liées à Fc et dans une moindre mesure des caractéristiques liées à la non-gaussianité sont détectées respectivement par CycloDetFcA et SK. Cela pourrait être utilisé comme une signature supplémentaire de la présence d'un RFI.

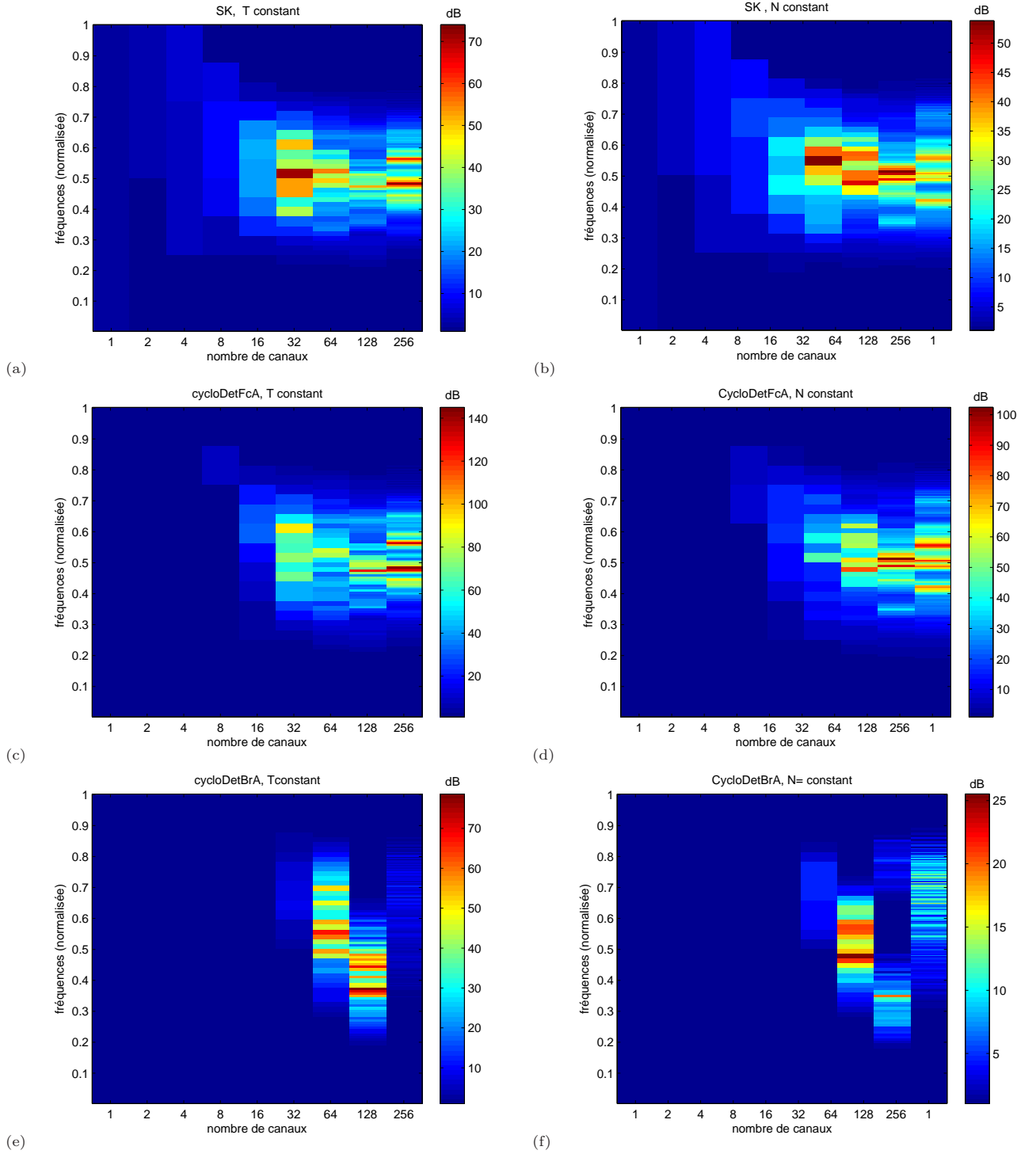


FIGURE 2.16 – Test de Fisher des détecteurs en forme de “carte” dans le cas du pulsar seul. $RBP = 10dB$. L’objectif est de tester la robustesse des détecteurs à la présence d’une impulsion pulsar forte.

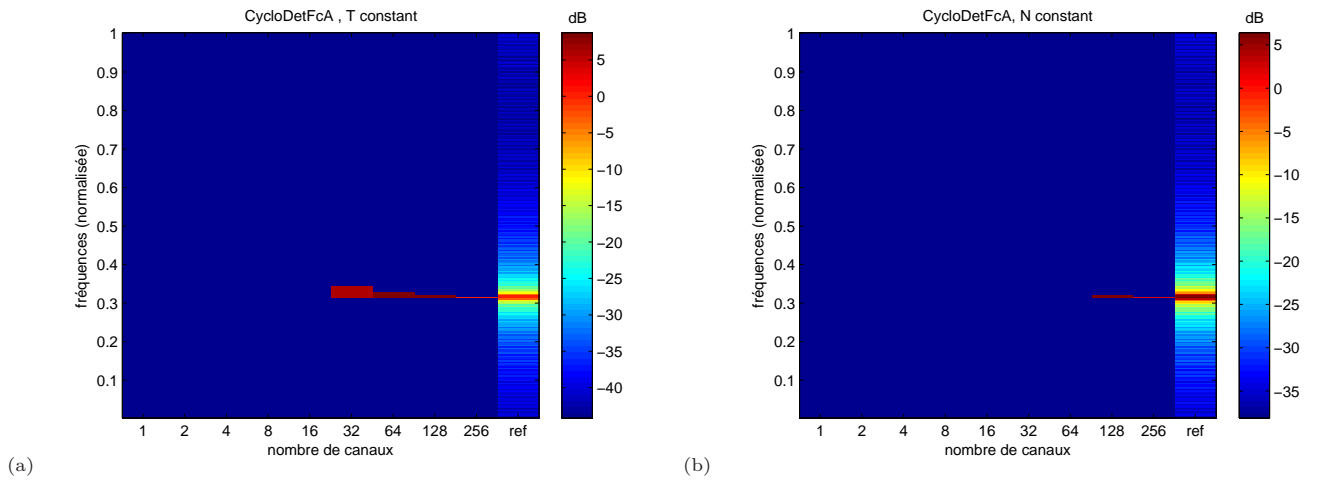


FIGURE 2.17 – Test de Fisher du détecteur CycloDetFcA en forme de “carte” en présence d’une interférence de type BPSK bande étroite ($Br = 1/100$) de très faible intensité ($RIB = -20dB$). Les autres paramètres de simulation sont $Fc = 0.31416$, $RBP = -20 dB$. Le brouilleur n’est détecté que lorsqu’il y a adéquation entre les caractéristiques spectrales de la RFI et la résolution spectrale du banc de filtres.

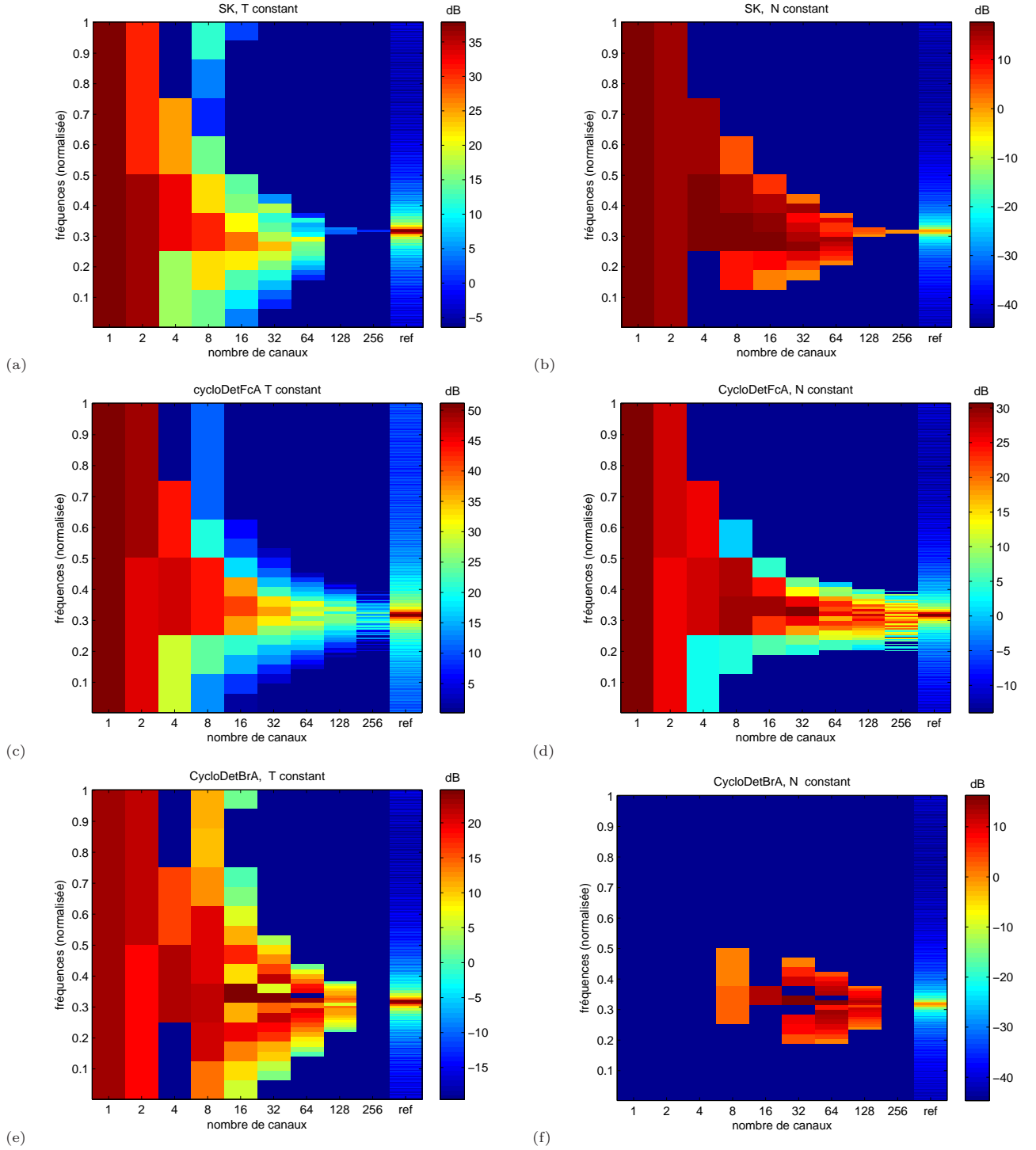


FIGURE 2.18 – Test de Fisher des détecteurs, cas de la RFI BPSK bande étroite : $Br = 1/100$ $Fc = 0.31416$ $INR = 10$ dB $PNR = -10$ dB.

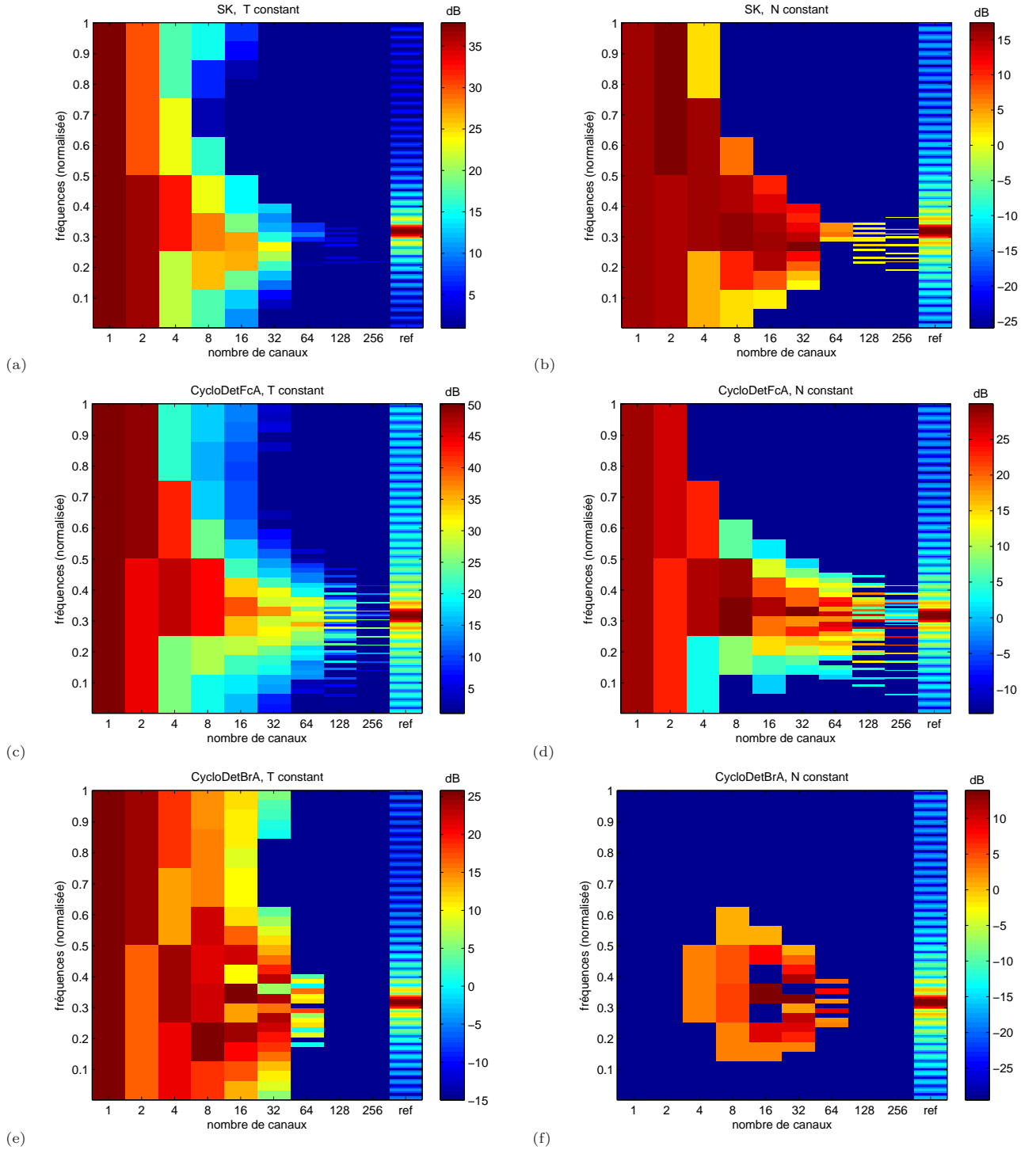


FIGURE 2.19 – Test de Fisher des détecteurs en forme de “map” en présence d’une interférence de type BPSK $Br = 1/37$ $Fc = 0.31416$ $RIB = 10$ dB $RPB = -10$ dB.

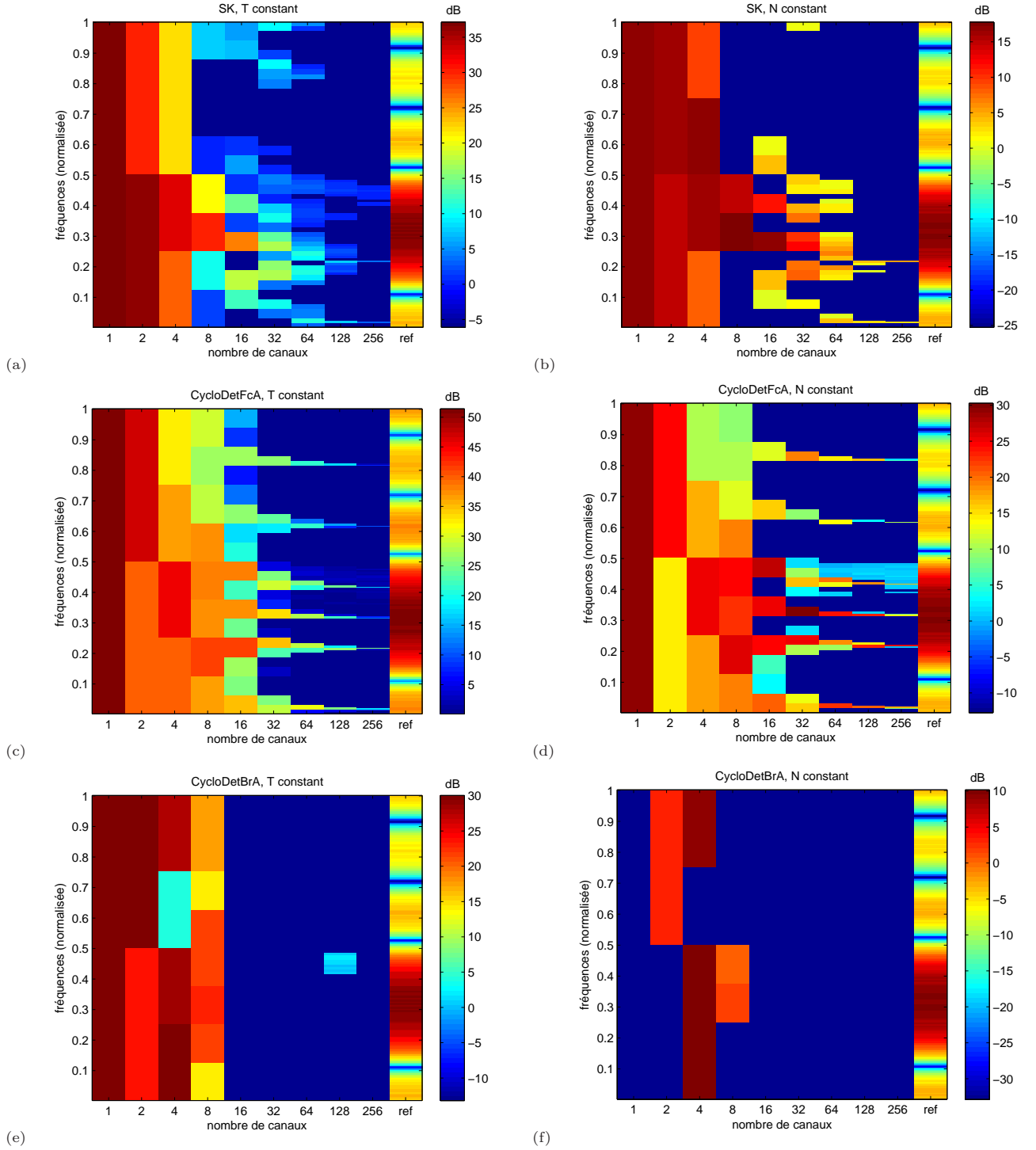


FIGURE 2.20 – Test de Fisher des détecteurs, cas de la RFI BPSK bande large : $Br = 1/5$ $Fc = 0.31416$
 $RIB = 10$ dB $RPB = -10$ dB

4. Conclusions :

Les résultats ci-dessus montrent que l'adéquation de la résolution temps-fréquence du récepteur et celle du brouilleur est importante. Si elle est possible, elle permet d'améliorer les performances des détecteur du fait de l'amélioration du *RIB* dans le canal fréquentiel considéré.

Ces résultats montrent également que l'utilisation de la cyclostationnarité peut être mis en défaut si la caractéristique temporelle (le baud rate par exemple) qui caractérise la cyclostationnarité mise en jeu est trop filtrée par le banc de filtre. La constatation est similaire pour la non-gaussianité.

Enfin, on voit que la gestion d'un ensemble de détecteurs à différents niveaux de résolution T-F pose le problème de la fusion de ces informations pour une exploitation en routine. Néanmoins, dans le cas de l'utilisation de la cyclostationnarité, les détecteurs aveugles qui sont proposés pourraient, en plus du niveau de détection qu'ils fournissent déjà, indiquer à quelle fréquence cyclique cette détection a eu lieu. La 2ème phase du projet Uniboard doit étudier plus en détail tout ces aspects. A la section suivante nous présentons la première phase de ce projet qui a vu l'implantation du détecteur CycloDetBrA.

2.8 Exemple d'implantation matériel : Uniboard

UNIBOARD est une carte électronique universelle et adaptée pour toute application en radioastronomie. Elle est développée par les instituts partenaires ASTRON et JIVE Néerlandais dans le cadre du projet européen RadioNet. La carte est constituée de huit composants programmables (FPGA)² de la société Altera, voir la figure 2.21(a).

Dans sa première phase de projet à la station de radioastronomie, UNIBOARD est dédiée à l'application "récepteur de signaux pulsar", la figure 2.21(b) montre l'ensemble des fonctions implantées qui constituent ce récepteur. Ce sont le banc de filtres polyphases, le dédisperseur, l'intégrateur synchrone (la fonction empilement), et les différents blocs liés à la détection et à la gestion des RFI. Les détecteurs des RFI qui ont été codés en VHDL et implantés par les ingénieurs de Nançay sont DetRadar et CycloDetBrA. Une fois que les RFI sont détectées, il faut choisir l'action appropriée pour garantir la qualité des observations obtenues. Deux options sont proposées :

- Soit fournir un signal (décision a) qui indique au récepteur la présence d'échantillons pollués. Ces échantillons sont alors mis de côté pour une analyse ultérieure. Ces échantillons peuvent même être classés par niveau de puissance. Cette option évite toute perte d'information pour le radioastronome. Il pourra décider en temps différé d'utiliser ou non les données détectées comme polluées.
- La seconde possibilité est plus radicale (décision b). Il s'agit de remplacer en temps réel les échantillons corrompus par des échantillons factices.

Les deux options ci-dessus peuvent être cumulées. Cela permet de conserver une trace des modifications imposées au signal et de donner le choix au radioastronome d'exploiter les deux versions des données, celle supposée nettoyée et celle uniquement identifiée comme polluée.

La nature des échantillons factices peut être définie suivant plusieurs possibilités :

- Échantillons simplement remplacés par des zéros ("blanking"). Son usage doit donc être limité aux cas où seulement quelques échantillons sont pollués, comme par exemple dans le cas du radar. Autrement, les statistiques des données risquent de changer.
- Échantillons choisis aléatoirement dans une zone propre du signal observé puis insérés aléatoirement à la place des échantillons pollués. Ceci est réalisé grâce à une mémoire circulaire contrôlée par le détecteur. Elle contient les N derniers échantillons non pollués de chaque canal.

Il faut signaler que ce protocole n'est pas propre à UNIBOARD. Il est applicable sur n'importe quelle instrumentation. En effet, il peut être aussi appliqué dans l'instrumentation pulsar aujourd'hui utilisée à la station de Nançay, qui est basée sur des GPUs. Dans ce cas, il faut réécrire les algorithmes en langage CUDA-C.

2. <http://minho.kasi.re.kr/SKA-Korea201008/Talks/SohnB.pdf>

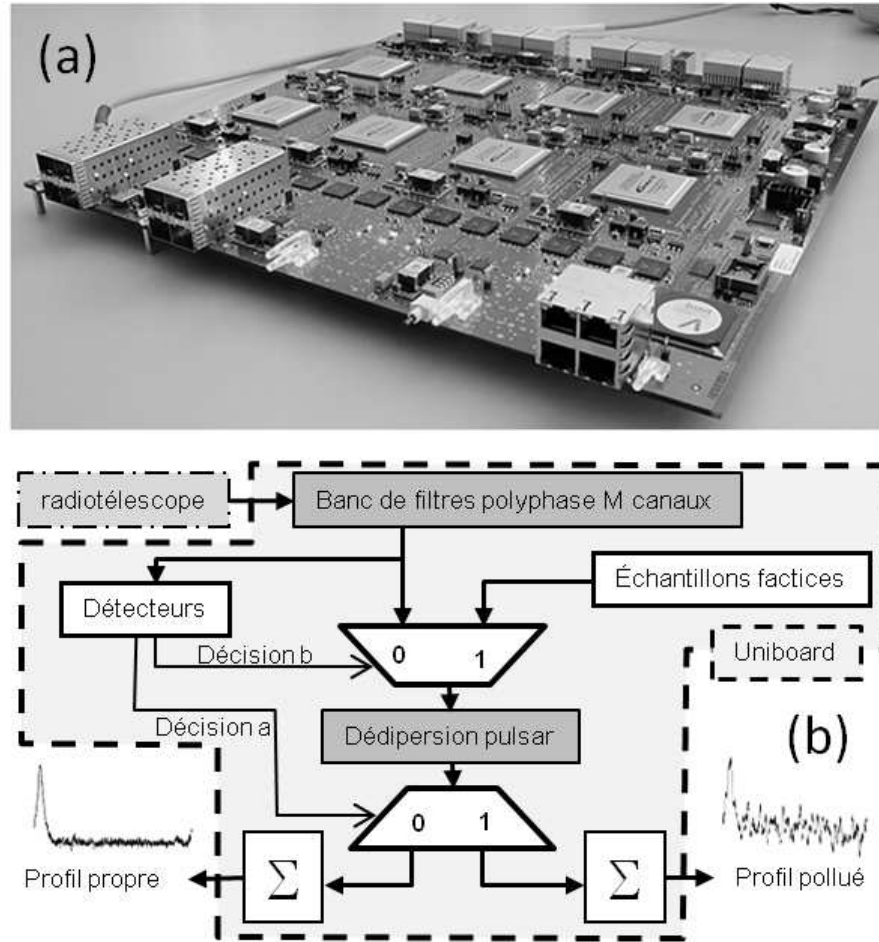


FIGURE 2.21 – le projet UNIBOARD. (a) : La carte Uniboard, développée par les instituts partenaires ASTRON et JIVE, est constituée de 8 composants programmables (FPGA) de la société Altera (Altera Stratix IV EP4SGX230KF40). Chaque FPGA possède 4 liaisons 10GbitE pour se connecter directement vers l'extérieur grâce à des interfaces à liaison optique ou cuivre, soit un débit total 160Gbit/s pour chaque colonne. Chaque FPGA dispose également de deux banques de mémoire DDR3 de 4Go. (b) Présentation des différentes fonctions implantées dans la carte dans le cadre de l'application " récepteur de signaux pulsar ". Cette application est réalisée en partenariat avec l'université de Manchester

2.9 Application sur des données réelles

LOFAR (LOW-Frequency ARray) est le précurseur d'une nouvelle génération de radiotélescopes. Conçu et construit par ASTRON, LOFAR est un grand réseau d'antennes basses fréquences (30-240 MHz) qui amène un gain de 1 à 2 ordres de grandeur en sensibilité et en résolution angulaire par rapport aux instruments existants. Si au départ le projet était purement néerlandais, il repose aujourd'hui sur un large partenariat européen. Il offre des bases interférométriques jusqu'à 100 km aux Pays-Bas et jusqu'à 1000 km en incluant les stations LOFAR européennes en Allemagne, au Royaume-Uni, en Suède, et en France à Nançay, voir la figure 2.22.

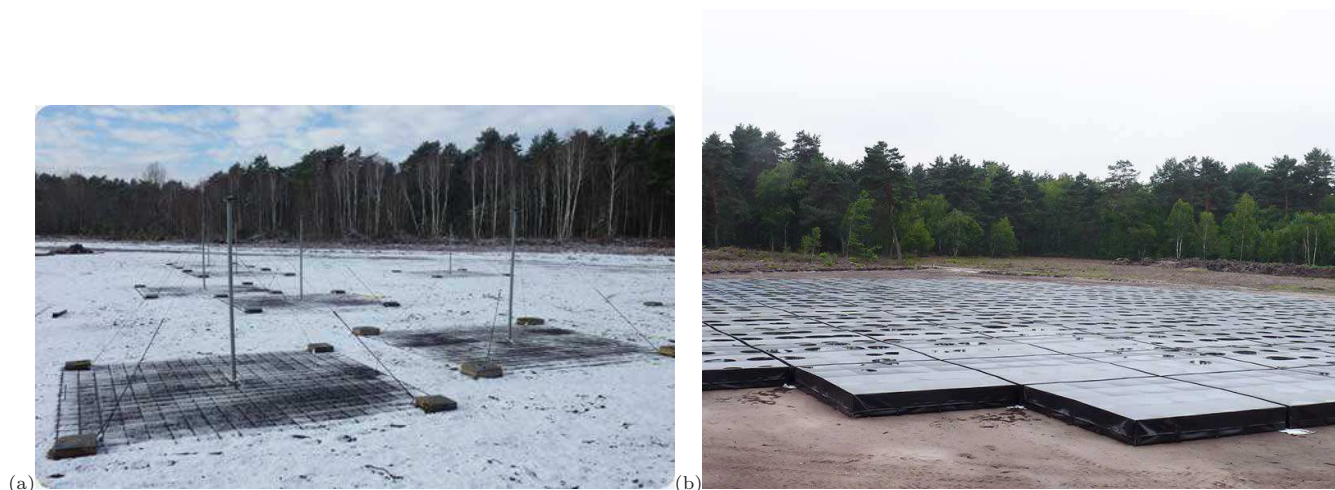


FIGURE 2.22 – LOFAR à la station de Nançay. (a) Les antennes basse fréquence (30-80 MHz) de LOFAR. (b) Les “tuiles” haute fréquence (110-250 MHz).

La construction de la station LOFAR en France à la station de radioastronomie à Nançay est achevée, et elle a commencé à observer en Mai 2011.

Nous avons testé, en temps différé avec le logiciel Matlab, l'algorithme CycloDetBrA sur des données enregistrées des premières observations faites par la station LOFAR de Nançay. Cette observation concerne le pulsar *PSR B0329+54* qui est à 2643 années lumière de la Terre dans la constellation de la Girafe. L'observation est faite dans la bande de fréquence $[142.9688, 149.0240]$ MHz, divisée en 31 sous-bandes de 32 canaux, et à la fréquence d'échantillonnage 6.104 KHz avec l'instrumentation Advanced Radio Transient Event Monitor and Identification System (ARTEMIS³). ARTEMIS est construit à base de GPU (Graphic Processing Unit).

La figure 2.23.(a) illustre un exemple d'un brouilleur bande étroite de grande puissance et de courte durée. On peut voir son influence sur la détection de l'impulsion individuelle après dédispersion dans la figure 2.23.(a).bas. La figure 2.23.(b) représente les résultats après l'application de CycloDetBrA pour chaque canal. La figure 2.23.(c) illustre un autre exemple de plusieurs brouilleurs à bande étroite, de petite puissance et continus. Dans ces cas, l'élimination des canaux brouillés augmente le rapport signal sur bruit de l'impulsion individuelle après dédispersion (exemple à la figure 2.23.(d)). Un autre exemple de brouilleur atypique d'une puissance énorme qui sature le plan T-F. CycloDetBrA élimine les canaux brouillés même avec une petite portion de ce brouilleur à cause de sa puissance. Bien qu'une bonne partie des canaux ait été masquée, les données propres sont récupérées avec une meilleure qualité qu'avant, voir la figure 2.23.(f).

Nous avons aussi testé le détecteur CycloDetBrA et SK sur une portion des données enregistrées par l'UTR-2 (The Ukrainian T-shaped Radio telescope, second modification). C'est le radiotélescope le plus grand dans les ondes radio décimétriques (voir figure 2.24). Cette partie de données ne contient pas d'impulsions pulsar mais représente un spectre qui n'est pas plat. Elle contient des brouilleurs à bande étroite et continus.

La figure 2.25 montre les résultats de l'application de CycloDetBrA et SK pour 2s de données issues d'un banc de filtres dans la bande de fréquence $[32.999737, 16.501049]$ MHz et échantillonnées à 4.028 KHz (voir la figure 2.25(a)). La figure 2.25(b) présente le plan T-F après élimination des canaux détectés brouillés par SK. La figure 2.25(c) présente le plan T-F après élimination des canaux détectés brouillés par CycloDetBrA. Remarquons que chaque détecteur a des "non-détection" à des canaux différents. On peut proposer alors la combinaison des deux détecteurs, vu la facilité de leur implémentation, pour indiquer les canaux brouillés. La

3. <http://www.oerc.ox.ac.uk/research/artemis>

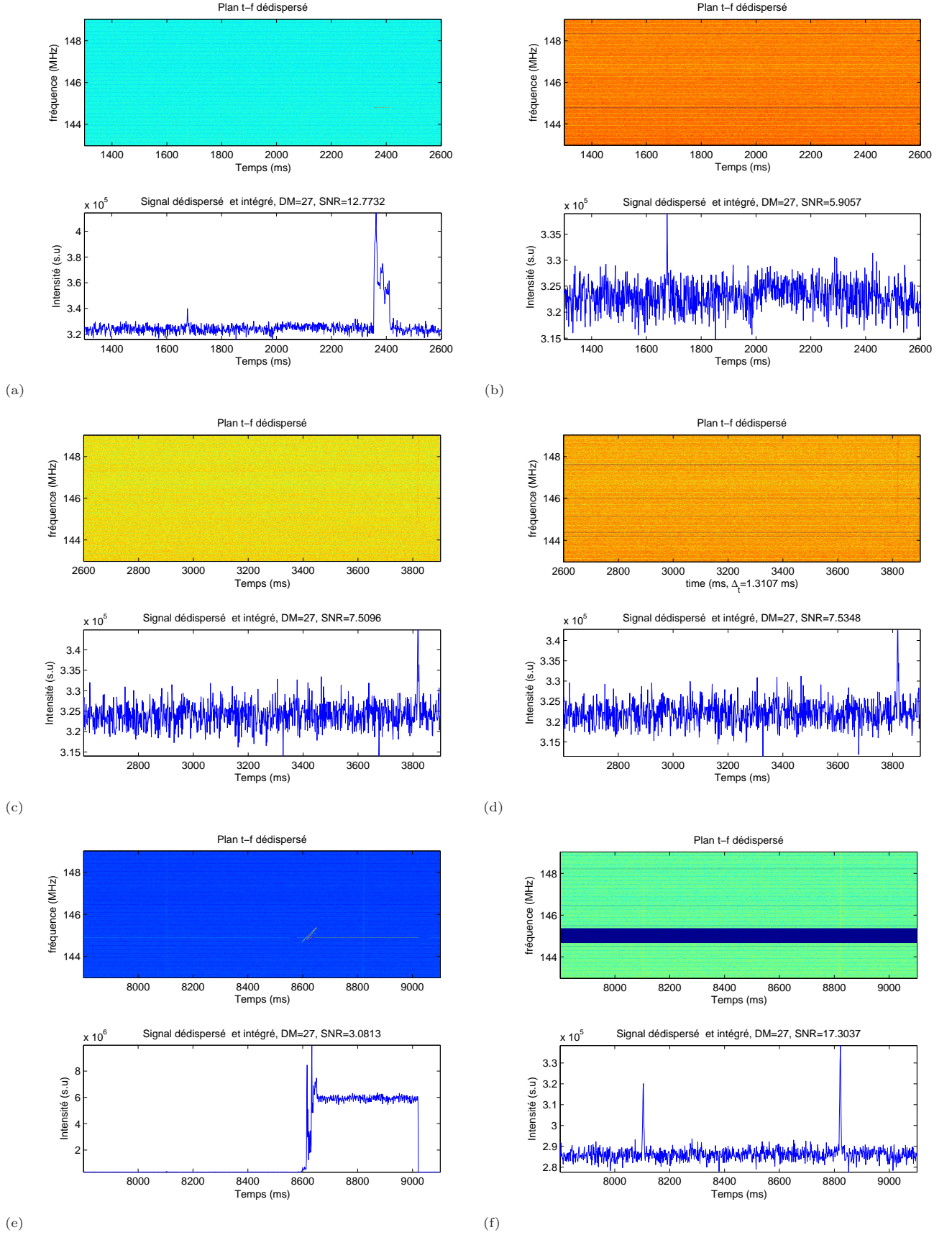


FIGURE 2.23 – Exemples d’application des détecteurs CycloDetBRA et SK pour l’observation du pulsar *B0329+54* avec LOFAR, 992 canaux fréquentiels, bande passante $BW = 6.0552\text{MHz}$ entre $[142.9688, 149.0240]\text{MHz}$, $\Delta f = 6.104\text{kHz}$ et $\Delta t = 163.84\mu\text{s}$. SNR représente le rapport de la valeur du maximum détecté sur l’écart-type du bruit.

figure 2.25(d) présente le plan T-F après élimination de tous les canaux indiqués comme brouillés par les deux détecteurs.



FIGURE 2.24 – Le radiotélescope UTR-2 (The Ukrainian T-shaped Radio telescope).

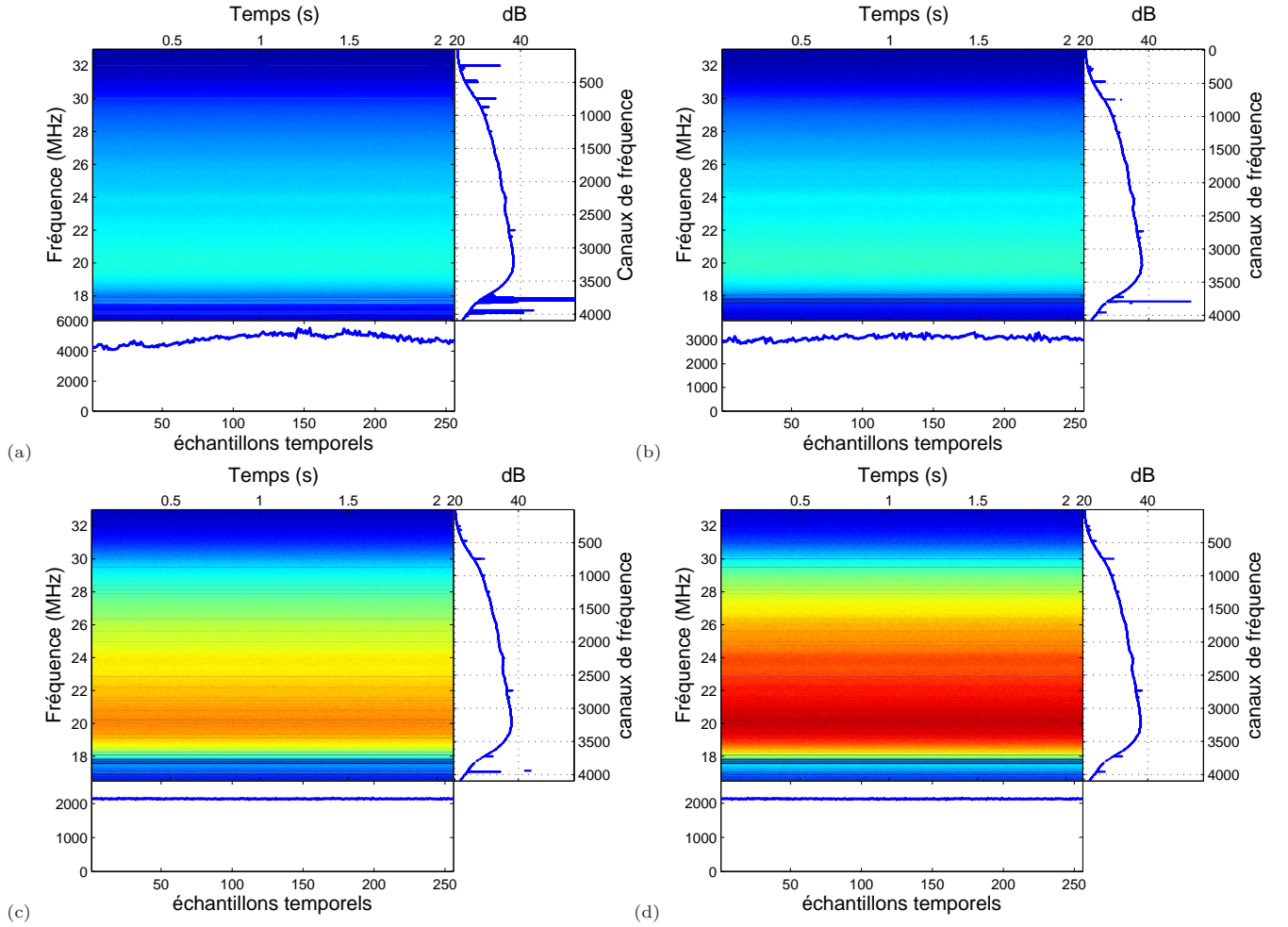


FIGURE 2.25 – Application sur les données de type "banc de filtres" pour les observations du pulsar *PSR B0031 – 07* : (a) le plan T-F de 2s des données pulsar *PSR B0031 – 07* enregistrées à la bande de fréquence $[32.999737, 16.501049]$ MHz et échantillonnées à 4.028KHz, $M = 4096$ canaux et $N = 256$ intants. (b) le plan T-F nettoyé en utilisant Kurtosis. (c) Plan T-F nettoyé en utilisant CycloDet. (d) Plan T-F nettoyé par les deux détecteurs.

2.10 Conclusion et perspectives

Dans le cadre du traitement des RFIs en radioastronomie, nous avons proposé une approche alternative aux techniques habituelles fondées sur des critères de puissance. Le critère proposé est fondé sur la cyclostationnarité potentielle des RFIs. Dans la cadre de l'observation des pulsars, nous avons montré que la cyclostationnarité était robuste face à la présence de pulses de pulsar et qu'elle offrait de meilleures performances par rapport à une approche fondée sur la non-gaussianité.

Dans l'optique d'une implantation sur une chaîne de réception, nous avons simulé les conséquences d'une approche permettant des tests sur plusieurs résolutions en fréquence. Cette approche permet d'optimiser les performances des détecteurs car elle tire mieux partie des caractéristiques fréquentielles de la RFI, toutefois il reste à définir une stratégie qui permette de coordonner l'ensemble des informations fourni par un tel système. Là encore la cyclostationnarité offre des perspectives par sa capacité à extraire du signal les signatures des RFIs, comme la fréquence de la porteuse ou le baud rate de la modulation. Il faut noter qu'en Télécommunication, des besoins similaires émergent via ce qu'on appelle la cognitive radio [Palicot, 2010]. Il pourrait être intéressant dans le futur de se plonger dans les approches proposées dans ce cadre, elles pourraient compléter celles que nous avons étudiées. En annexe A, une étude bibliographique rapide est proposée.

Dans cette section, nous avons également évoqué via l'application UNIBOARD une stratégie possible de traitement des données détectées comme polluées. Outre l'action classique d'éliminer les données, ou de les remplacer par des données factices, il nous semble important de pouvoir juste étiquetter les données en fonction

de leur qualité. Cet objectif est d'autant plus important que la nouvelle génération de radiotélescopes (LOFAR, SKA) vont générer des volumes des données très important. A l'astronome ensuite de choisir ce qu'il va exploiter.

En d'autres termes :

- il faut se servir des détecteurs uniquement pour qualifier la qualité des données.
- il faut amener cette information à des nœuds de décisions plus stratégiques. Possibilité de fusion de données pour prendre des décisions plus adaptées. Possibilité de segmenter les résultats finaux d'observations en fonction de la qualité des données qui ont servi à calculer ces résultats.
- il faut tracer cette information (qualité des données et décisions/actions prises) dans une base d'informations. La donnée est alors identifiée sur la base des grilles d'analyse dont il est question ci-dessus.

L'ensemble de ces fonctions doit se faire en temps réel, mais cette base d'information doit également être accessible en temps différé lors du dépouillement des résultats.

Chapitre 3

Détection d'événements impulsionsnels intermittents

3.1 Introduction

Les observations radio-astronomiques des pulsars peuvent être classées en deux modes ; le mode recherche de nouveaux pulsars "Survey" et le mode de chronométrage "Timing" comme cela a été abordé au chapitre 1. Le mode "Survey", qui nous intéresse ici, peut être mis en oeuvre de deux façons différentes. La première méthode commence par une dédispersion pour un ensemble de valeurs de DM, suivi d'une recherche de périodicité après une transformée de Fourier, dont l'objectif est d'avoir une liste de couples (DM,période) susceptibles d'être de nouveaux pulsars. Le deuxième usage du mode "Survey" nécessite également une dédispersion pour un ensemble de plusieurs valeurs de DM, puis une inspection des séries temporelles dédispersées à la recherche d'impulsions individuelles "Single pulses". Dans ce cas, les détections peuvent être des impulsions pulsars assez fortes pour être détectables individuellement, comme elles peuvent être des RFIs ou des variations imprévues dans le bruit. Les impulsions individuelles pulsars peuvent être ce qu'on appelle les impulsions géantes "giant pulses" ou bien l'impulsion de pulsars dont nous ne recevons qu'une très faible fraction des impulsions, appelés RRATs ("Rotating RAdio Transients").

Par définition, une impulsion géante est une impulsion intense et étroite. Typiquement, elle est d'énergie bien plus forte que celle d'un pulse moyen et elle apparaît au voisinage de la même phase rotationnelle que l'impulsion principale. Les impulsions géantes sont observées dans deux catégories de pulsars, les jeunes énergétiques pulsars comme le Crabe, qui a été d'ailleurs découvert grâce à son émission d'impulsions géantes [Staelin and Reifenstein, 1968], et dans les pulsars millisecondes. Les pulsars émetteurs d'impulsions géantes ont un champ magnétique rapporté au niveau du cylindre lumière (là où la co-rotation se ferait à la vitesse la lumière) très fort comparé à celui des pulsars qui ne les émettent pas [Becker, 2009].

Les RRATs sont aussi des étoiles à neutrons en rotation dont l'émission est vue comme une suite d'impulsions isolées mais avec une distribution de puissance différente de celles émis par un pulsar normal. En général, les RRATs ne sont pas détectables par la recherche de périodicité et la plupart de ceux détectés ont des périodes assez longues, par exemple 5/10 ont une période supérieure à 4s [McLaughlin et al., 2006], alors que ce n'est le cas que d'un peu plus d'une cinquantaine sur l'ensemble des ~ 2000 pulsars ordinaires connus.

La recherche en "mode impulsion individuelle" est importante pour la découverte de ces pulsars qui ont de telles propriétés d'émission. La plupart des pulsars "normaux" sont détectables avec un grand rapport signal sur bruit en mode recherche de périodicité. Alors que les pulsars avec des impulsions individuelles occasionnelles ou les pulsars avec des périodes longues sont détectables avec un grand rapport signal sur bruit en "mode impulsion individuelle". Un exemple qui illustre l'importance du "mode impulsion individuelle" est le pulsar *PSR J1918+18* découvert avec la recherche d'impulsions individuelles faite par Arecibo [Nice, 1999] mais qui n'était pas détecté avec la recherche de périodicité [Nice and Taylor, 1995].

L'étude faite dans ce chapitre se place dans le cadre de la recherche d'événements impulsionsnels intermittents correspondant aux deux types décrits ci-dessus.

3.1.1 Détection d'impulsions individuelles avec PRESTO

Le logiciel le plus utilisé pour rechercher des candidats pulsars à partir de séries temporelles est PRESTO (PulsaR Exploration and Search TOolkit) écrit par Scott Ransom¹. Ce logiciel effectue plusieurs fonctions nécessaires au traitement des données pulsars : dédispersion, recherche dans le domaine de Fourier (recherche de périodicité), recherche d'impulsions individuelles, empilement, en plus d'autres fonctions comme la recherche et l'élimination de parasites par exemple.

Le traitement pour la détection d'impulsions individuelles se découpe en plusieurs étapes qu'on résume ci-dessous :

- **Recherche des interférences radio fréquence** Le programme calcule des statistiques du signal reçu dans les domaines temporel fréquentiel à la recherche des RFIs. Les zones détectées sont ensuite remplacées par la moyenne du bruit [Lorimer and Kramer, 2005].
- **Dédispersion** La dédispersion de n échantillons se fait pour différentes valeurs de la mesure de dispersion DM appartenant à une liste prédéfinie. Le choix de l'écart entre deux valeurs de DM successives est très important. Il ne doit pas être trop grand pour ne pas étaler l'impulsion, ni trop petit pour minimiser le temps de calcul. Pour le traitement des données du programme de recherche 'Survey FOSTER'² à Nançay, les valeurs du DM à utiliser pour la dé-dispersion sont entre $[0, 1500]\text{pc.cm}^{-3}$. Produite par un script de PRESTO (DDplan.py), la figure 3.1 montre le pas optimum (DM Stepsize Smearing en rouge) entre deux valeurs de DM successives. Les différents effets produisant l'élargissement de l'impulsion pour un DM donné sont calculés en fonction du pas de l'échantillonnage dt , de la bande passante BW , de la fréquence centrale f_{ctr} et du nombre de canaux N_{chan} .
- **Recherche Single pulse** Une fois que la série temporelle est dédispersée pour un DM donné, PRESTO cherche des événements transitoires supérieurs à un seuil fixé. PRESTO a été utilisé pour la détection des événements intermittents dans le cadre du retraitement des observations FOSTER [Desvignes, 2009] et le seuil a été fixé à 5σ .

Nous proposons ici une méthode innovante pour la détection des impulsions individuelles sans avoir à dédispenser la série temporelle pour de nombreuses valeurs de DM, tout en étant robuste contre les RFIs qui sont soit intermittentes à large bande, soit continues à bande étroite. Cette méthode est basée sur la transformée de Fourier à deux dimensions et la transformée de Radon, et elle est applicable en temps réel.

Dans la section suivante, nous expliquons cette nouvelle méthode intitulée SIPSFAR (SIngle Pulse Search using Fourier And Radon transforms). Ensuite, nous détaillons l'étude comparative de la sensibilité des deux méthodes qui sera suivie par la présentation des résultats des simulations ainsi que des résultats obtenus en testant l'algorithme sur des données type 'survey' du pulsar du Crabe (N_p canaux de fréquences échantillonnés tous les dt).

3.2 Détection de pulses individuelles avec SIPSFAR

Les nouvelles générations des radiotélescopes avec plus de sensibilité, plus de résolution angulaire, dans des bandes de fréquences plus larges, ont suscité un nouveau besoin de traitement des données en temps réel, afin de minimiser le volume de données à stocker. Dans cette optique, nous proposons un nouvel algorithme de recherche d'événements impulsionnels intermittents (recherche de single pulse) capable de fonctionner en temps réel et en aveugle. Ici, le terme aveugle désigne le fait que l'algorithme peut détecter une impulsion individuelle sans dédispenser au préalable à une valeur de mesure de dispersion DM donnée. L'algorithme SIPSFAR est également robuste contre certains types de RFIs ; les RFIs impulsionnels à large bande et les RFIs continus à bande étroite.

3.2.1 Principe de SIPSFAR

La méthode SIPSFAR est basée sur le fait que dans un plan temps-fréquence, une impulsion pulsar peut être représentée par une droite. Cette hypothèse est valide si le couple (DM, bande d'analyse) permet de négliger la courbure du retard dispersif. Ainsi, on commencera à s'écarter de l'hypothèse d'une dispersion linéaire si le pulsar présente une grande valeur de mesure de dispersion DM et/ou si la bande de fréquence analysée est grande.

1. <http://www.cv.nrao.edu/~sransom/presto/>

2. un programme de recherche de nouveaux pulsars, cf. 4.2

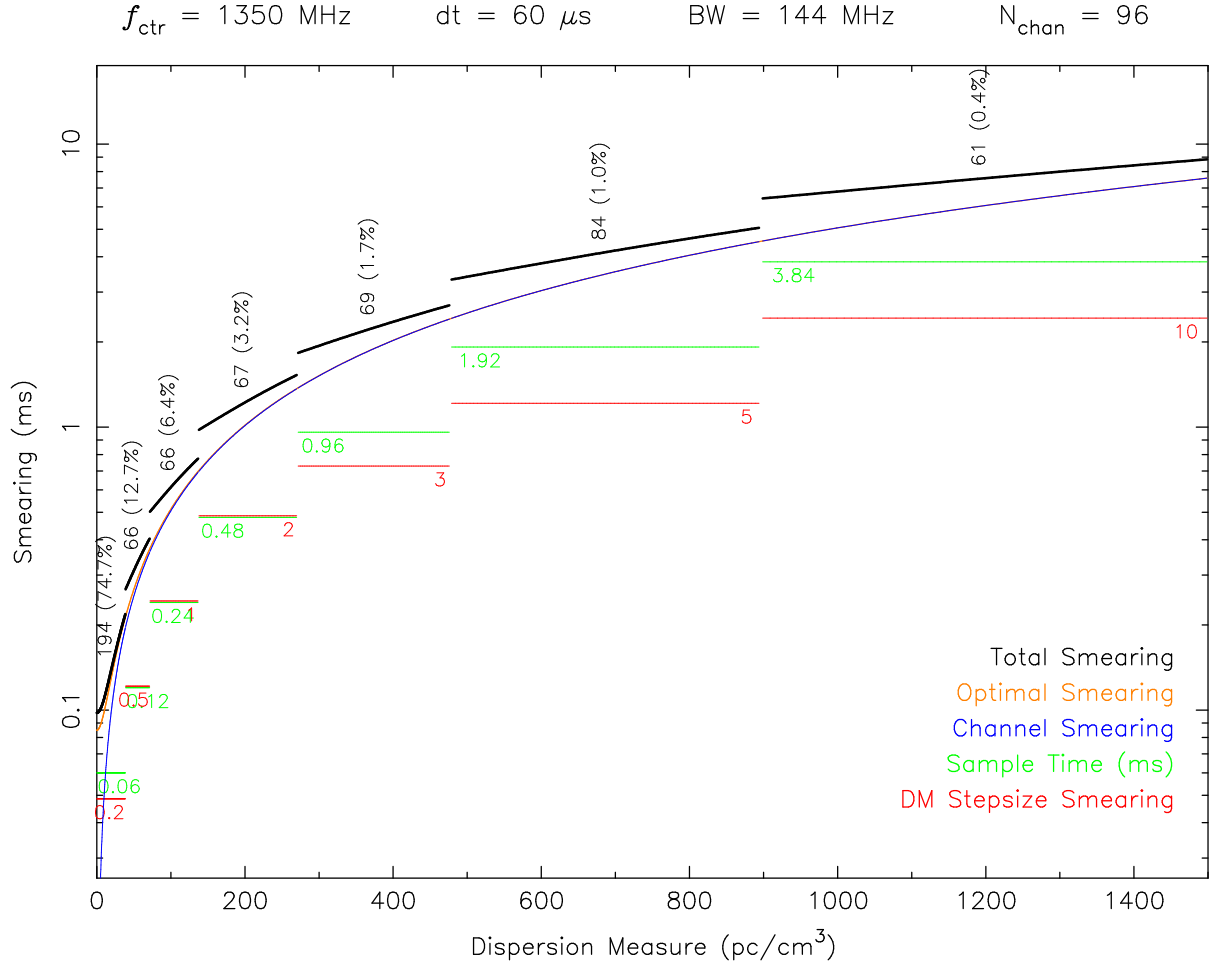


FIGURE 3.1 – Plan de dédispersion pour les données NBPP montrant l'étalement temporel de l'impulsions (en ms) en fonction de la mesure de dispersion DM et en prenant en compte toutes les contributions (élargissement produit par l'échantillonnage temporel et par la largeur du canal de fréquence). La courbe en rouge donne le pas optimum du DM (en $\text{pc}.\text{cm}^{-3}$, et la courbe en vert donne le pas adéquat de l'échantillonnage dt en ms .

L'algorithme SIPSFAR, qui est donc basé sur l'hypothèse d'une dispersion linéaire dans la bande de fréquence, utilise une transformée de Radon (voir la section suivante 3.2.2 et l'annexe C) qui sera appliquée sur la transformée de Fourier 2D du plan T-F. La transformée de Radon, en additionnant l'intensité de la transformée de Fourier le long de lignes radiales, concentre le signal pour la ligne où l'impulsion du pulsar a été projetée par la FFT.

L'algorithme peut être divisé en plusieurs étapes. La figure 3.2 représente un organigramme expliquant l'algorithme :

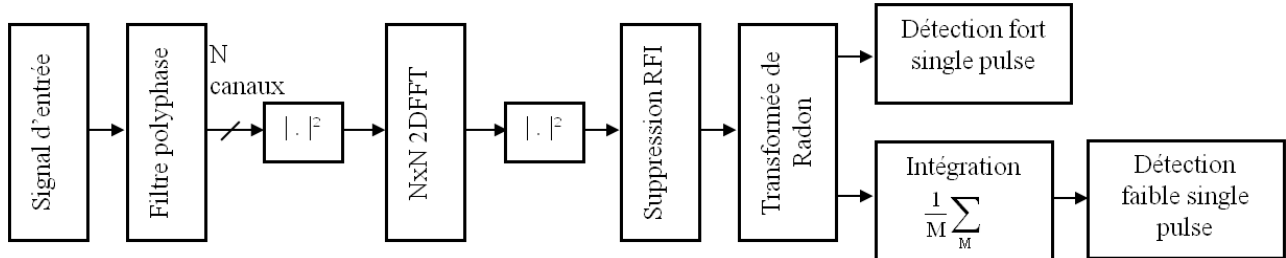


FIGURE 3.2 – Description de l'algorithme SIPSFAR

- Après la conversion analogique-numérique, un plan temps-fréquence (t-f) de puissance de la forme d'onde du signal est produit par un banc de filtres numériques (voir un exemple dans la figure 3.3(a)). Il s'agit d'une fonctionnalité en temps réel commune à tous les radiotélescopes.
- En temps réel des transformées de Fourier en deux dimensions ($2D - FFT$) sont appliquées sur des images T-F successives de dimensions $N \times N$. Un exemple est donné dans la première ligne de 3.3(b). Avec cette approche, toutes les impulsions d'un pulsar dispersé sont projetées au même endroit sous la forme d'une ligne de l'image $2D - FFT$. L'angle de cette ligne dépend uniquement de la mesure de dispersion DM du pulsar et est différent de celui obtenu par les RFIs. En effet, toute RFI impulsive à large bande (respectivement RFI continue à bande étroite) sera projetée sur l'orthogonal de sa direction, c-à-d. la ligne horizontale (respectivement la verticale) au centre de l'image $2D - FFT$. Ainsi, pour retirer les RFIs et obtenir une image propre, on élimine la croix centrale de l'image $2D - FFT$.
- Une transformée de Radon est appliquée sur l'image $2D - FFT$ nettoyée. Elle consiste à additionner l'intensité d'image le long de lignes radiales successives. Lorsque cette ligne d'intégration correspond à la ligne où l'impulsion du pulsar a été projetée, toute la puissance de l'impulsion est concentrée en un angle. Pour les autres angles d'intégration, seule la contribution du bruit va être intégrée. La deuxième ligne de 3.3(b) illustre ce principe.

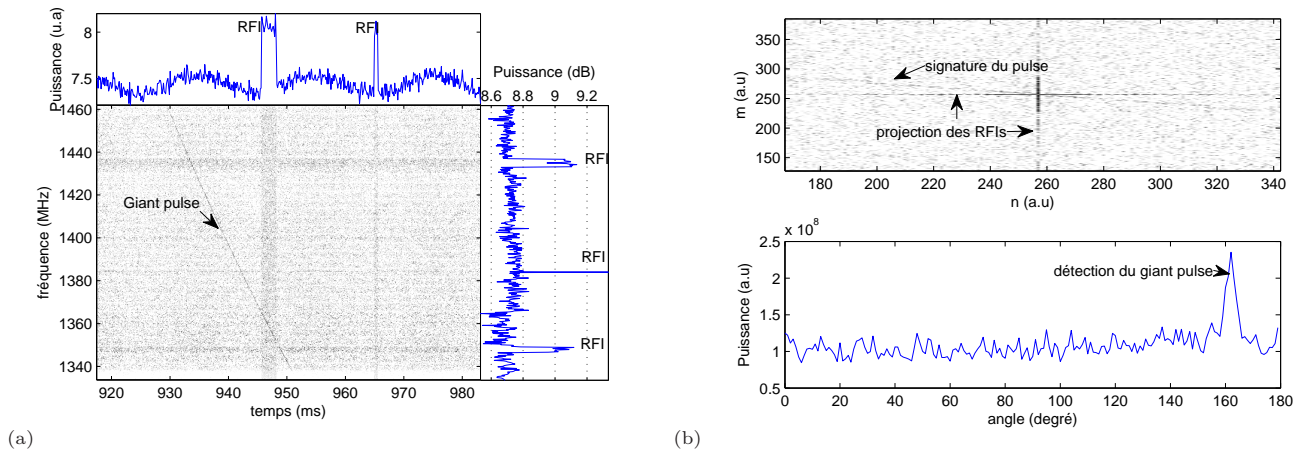


FIGURE 3.3 – Test de transformée de Radon sur un bloc de données contenant une impulsion géante du pulsar du Crabe

L'avantage de cette approche est que seule une quantité limitée de données est stockée. Si le résultat de la transformée de Radon dépasse un seuil prédéfini auparavant, et comme les RFIs sont éliminés, alors le profil Radon représente probablement la détection d'une impulsion individuelle. Il est stocké afin de pouvoir retrouver

le DM ultérieurement. En outre, il fournit des capacités simples d'atténuation des RFIs. Les profils de Radon successifs peuvent également être intégrés au fil du temps (pour M images de $N \times N$ $2D-FFT$) afin d'améliorer la détection des impulsions. Dans ce qui suit, nous étudions et comparons la sensibilité de SIPSFAR et celle de la méthode classique souvent utilisée dans PRESTO (voir section 3.1.1).

3.2.2 La transformée de Radon

1. Pourquoi la transformée de Radon

La figure 3.4, pour illustration, montre que la transformée de Radon (TR) est bien adaptée pour la détection des lignes dans une image. La figure 3.4(a) présente une image avec quatre lignes noyées dans un bruit gaussien. La transformée de Radon de cette image est illustrée par la figure 3.4(b), on remarque la présence de quatre points lumineux ayant des paramètres différents. Il s'agit de la détection des quatre lignes de l'image. On remarque que chaque ligne dans l'image a son couple de paramètres (ρ, θ) . ρ est la position par rapport au centre de l'image et θ est l'angle d'inclinaison par rapport à l'horizontale. En effet, la transformée de Radon est capable de transformer chaque ligne en un maximum positionné en fonction des paramètres de la ligne. Ainsi, une problématique de détection de ligne dans une image brouillée se transforme en une décision de seuil dans l'image obtenue par la transformée de Radon.

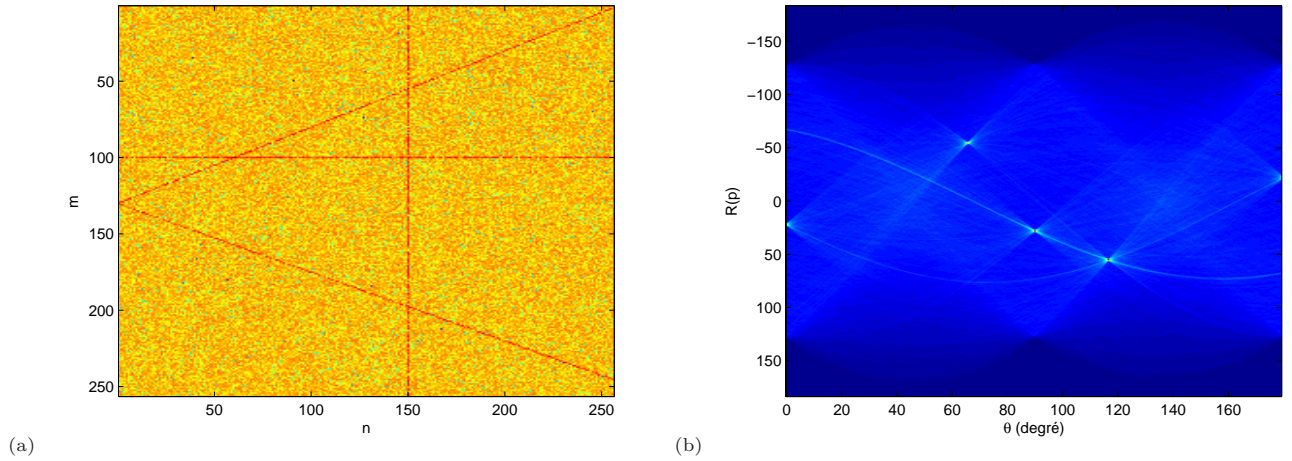


FIGURE 3.4 – Illustration de l'intérêt de la transformation de Radon : (a) une image de taille $m \times n$ où quatre lignes sont noyées dans un bruit. (b) la TR de l'image précédente, le premier maximum à $\theta = 0^\circ$ représente la droite horizontale. Celui à 90° représente la droite verticale. Le maximum à 60° représente la droite ayant la pente positive et le dernier à 120° représente celle ayant la pente négative.

2. La transformée de Radon discrète

La transformée de Radon discrète est définie pour une image carrée, où les échantillons repérés par leurs coordonnées (x, y) sont symétriques par rapport au centre de l'image :

$$x_{min} = -x_{max} = -\frac{M-1}{2}\Delta x \quad (3.1)$$

$$y_{min} = -y_{max} = -\frac{M-1}{2}\Delta x \quad (3.2)$$

$$\rho_{min} = -\rho_{max} = -\frac{K-1}{2}\Delta \rho \quad (3.3)$$

La transformée de Radon discrète pour une fonction continue $g(x, y)$ est donnée par (davantage de détail dans l'annexe C) :

$$\check{g}(\rho, \theta) \approx \begin{cases} \frac{\Delta x}{|\sin\theta|} \sum_{m=0}^{M-1} g(m, \left[-m \cot\theta + \frac{\rho - x_{\min}(\cos\theta + \sin\theta)}{\Delta x \sin\theta}\right]) & \text{si } \sin\theta > \frac{1}{\sqrt{2}} \\ \frac{\Delta x}{|\cos\theta|} \sum_{n=0}^{M-1} g(-n \tan\theta + \left[\frac{\rho - x_{\min}(\cos\theta + \sin\theta)}{\Delta x \cos\theta}\right], n) & \text{si } \sin\theta \leq \frac{1}{\sqrt{2}} \end{cases} \quad (3.4)$$

où $M = N$, $\Delta x = \Delta y$, et $x_{\min} = y_{\min}$.

3. La transformée de Radon ($\rho = 0, \theta$)

Pour minimiser le nombre d'opérations à effectuer, nous avons considéré la DRT (Discrete Radon Transform) pour $\rho = 0$. En tenant compte de l'équation de l'annexe C.10 et $\Delta x = 1$ l'équation C.14 devient :

$$\check{g}(0, \theta) \approx \begin{cases} \frac{1}{|\sin\theta|} \sum_{m=0}^{M-1} g(m, \left[(-m + \frac{M-1}{2}) \cot\theta + \frac{M-1}{2}\right]) & \text{si } \sin\theta > \frac{1}{\sqrt{2}} \\ \frac{1}{|\cos\theta|} \sum_{n=0}^{M-1} g\left(\left[(-n + \frac{M-1}{2}) \tan\theta + \frac{M-1}{2}\right], n\right) & \text{si } \sin\theta \leq \frac{1}{\sqrt{2}} \end{cases} \quad (3.5)$$

La figure 3.5 montre que pour une droite qui passe par le centre de l'image (l'origine de la DRT), l'application de la DRT à la position $\rho = 0$ détecte la droite noyée dans le bruit à l'angle $\theta = 135^\circ$.

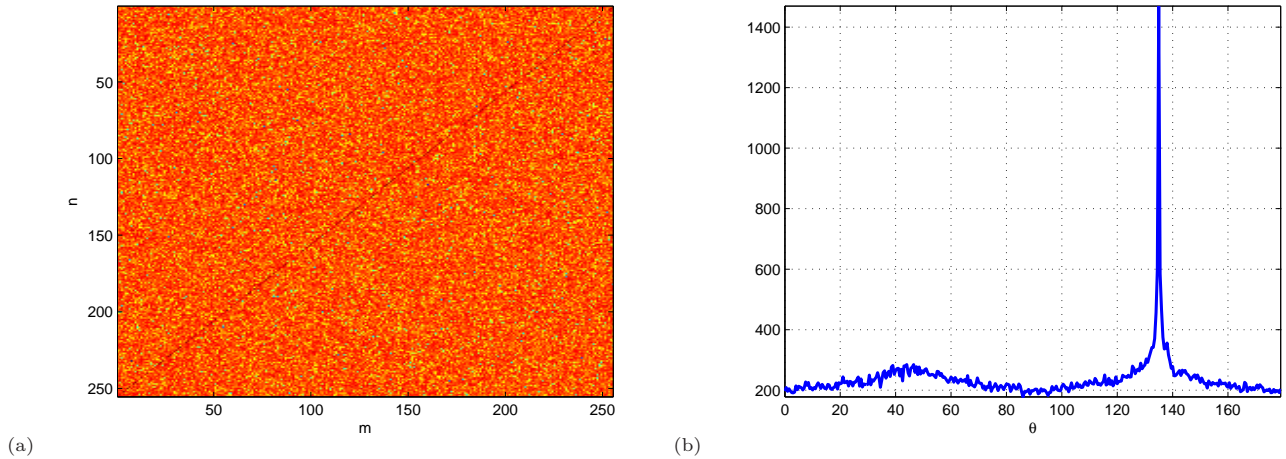


FIGURE 3.5 – Illustration de la Discrete Radon Transform (DRT) pour $\rho = 0$: (a) une image avec une droite passant par le centre ayant une inclinaison par rapport à l'horizontale égale à 135° à partir de la gauche de l'image. (b) la DRT($\rho = 0, \theta$) appliquée à l'image avec une détection à $\theta = 135^\circ$

Dans l'algorithme SIPSFAR, c'est la transformée de Radon discrète définie par l'équation 3.5 qui est utilisée.

3.2.3 Comparaison de la sensibilité

On considère le modèle du signal et les paramètres suivants pour le calcul des performances des deux méthodes :

- Le bruit électronique est un bruit blanc gaussien complexe centré de puissance σ_b^2 ; $\mathcal{N}_{\mathbb{C}}(0, \sigma_b^2)$.
- Le signal du pulsar est un signal aléatoire gaussien complexe centré de puissance σ_p^2 $\mathcal{N}_{\mathbb{C}}(0, \sigma_p^2)$. Pour des raisons de simplifications, on considère la forme de l'impulsion rectangulaire de durée L . Le rapport signal sur bruit est défini par $SNR = \sigma_p^2 / \sigma_b^2$. Pour le calcul théorique de la sensibilité, on suppose $\sigma_p^2 \ll \sigma_b^2$.
- Il n'y a pas de RFI.
- le plan temps-fréquence de puissance est de taille $N \times N$. Les axes du temps et la fréquence sont normalisés.

- Pour l'algorithme PRESTO, on ne considère que l'étape où la mesure du DM est correcte et où l'impulsion a été dé-dispersée avec. Ainsi, on peut supposer que :
 1. $DM = 0$, c'est à dire une impulsion verticale dans le plan T-F de puissance.
 2. Nous supposons que la re-synchronisation de l'impulsion par rapport à la période de rotation du pulsar a été faite. Les impulsions se localisent alors toujours dans la même position dans le plan t-f.
 3. Le plan t-f est intégré le long des fréquences pour donner un profil individuel de l'impulsion.
 4. Finalement, M impulsions individuelles sont sommées. La sensibilité S_{PRESTO} est définie par le rapport de l'amplitude de l'impulsion dans le profil intégré sur l'écart type du bruit, voir figure 3.6.(b).
- Pour SIPSFAR, on suppose aussi que $DM = 0$. En effet, pour simplifier les calculs, on considère que $DM = 0$ est juste une valeur de DM parmi les autres. La croix centrale de l'image $2D - FFT$ est systématiquement éliminée. La moyenne de M profils Radon fournit le profil intégré nettoyé. La sensibilité $S_{SIPSFAR}$ est définie par le rapport de l'amplitude du profil Radon à l'angle $\theta = 90^\circ$ ($DM = 0$) sur l'écart type estimé à partir des autres valeurs de θ .

Ce modèle est illustré dans la figure 3.6(a). Une définition graphique de la sensibilité est donnée par la figure 3.6(b).

Après avoir énuméré les hypothèses ci-dessus, les sensibilités des deux méthodes sont calculées comme suit (le détail du calcul est dans l'annexe D).

La sensibilité de la méthode SIPSFAR est :

$$S_{SIPSFAR} = T^2 \sqrt{NMSNR^2} \quad (3.6)$$

Et la sensibilité de la méthode PRESTO est :

$$S_{PRESTO} = \sqrt{NMSNR} \quad (3.7)$$

Dans la figure 3.6(c), ces sensibilités ont été calculées par simulation pour différentes valeurs des paramètres. Les sensibilités théoriques sont aussi tracées. A partir de cette figure, on peut remarquer ce qui suit :

- Les mesures des sensibilités avec les simulations coïncident avec les valeurs théoriques attendues.
- La sensibilité de la méthode SIPSFAR est proportionnelle à SNR^2 . Ceci rend cette approche moins sensible pour les faibles SNR comparé à la méthode PRESTO, voir par exemple les courbes pour $SNR = -5dB$, $T = 1$ dans la figure 3.6. Ce point faible est balancé par le fait que SIPSFAR est sensible aussi à la durée T de l'impulsion dans le plan t-f. Plus l'impulsion est large, mieux est la détection avec SIPSFAR, voir la cas de $SNR = -5dB$, $T = 8$.
- Ces performances sont atteintes sans une connaissance a priori du pulsar (ni DM ni période). Ceci n'est pas le cas pour la méthode PRESTO, où ces performances sont uniquement atteintes si la valeur de DM est connue et exacte.

Avec la méthode SIPSFAR, on peut avoir une première estimation de la mesure du DM. La mesure de dispersion est fonction de l'angle de l'inclinaison α de l'impulsion par rapport à la verticale dans le plan temps fréquence, elle est donc fonction de l'angle θ par rapport à l'horizontale dans l'image de la $2D - FFT$ pour lequel Radon fait la détection. Pour des DM grands, sommer des échantillons temporels successifs dans chaque canal du plan T-F est la solution pour parvenir à contenir l'impulsion dans un seul plan T-F. Le nombre d'échantillons sommés n_{bin} nécessaire varie selon la valeur du DM. L'estimation du DM avec la méthode SIPSFAR est calculée à partir de la relation donnant l'angle de Radon :

$$\tan(\alpha) = \frac{t_{ech}}{N_{chan}} \quad (3.8)$$

avec t_{ech} la durée de l'impulsion dans le plan t-f en nombre d'échantillons et N_{chan} la bande de fréquence totale en nombre de canaux. Rappelons que le plan t-f est carré ($t_{ech} = N_{chan}$). Notons que $\alpha = \theta$. Le retard de dispersion t_{DM} est :

$$t_{DM} = t_{ech} \times n_{bin} \times dt = t_{ech} \times n_{bin} \times dt \times \tan(\theta) \quad (3.9)$$

avec dt la durée d'échantillonnage. Finalement, l'équation 3.9 et l'équation 1.4 donnent :

$$DM = \frac{N_{chan} \times f_0^3 \times dt \times n_{bin}}{8.3 \times 10^6 \times BW} \tan(\theta) \quad (3.10)$$

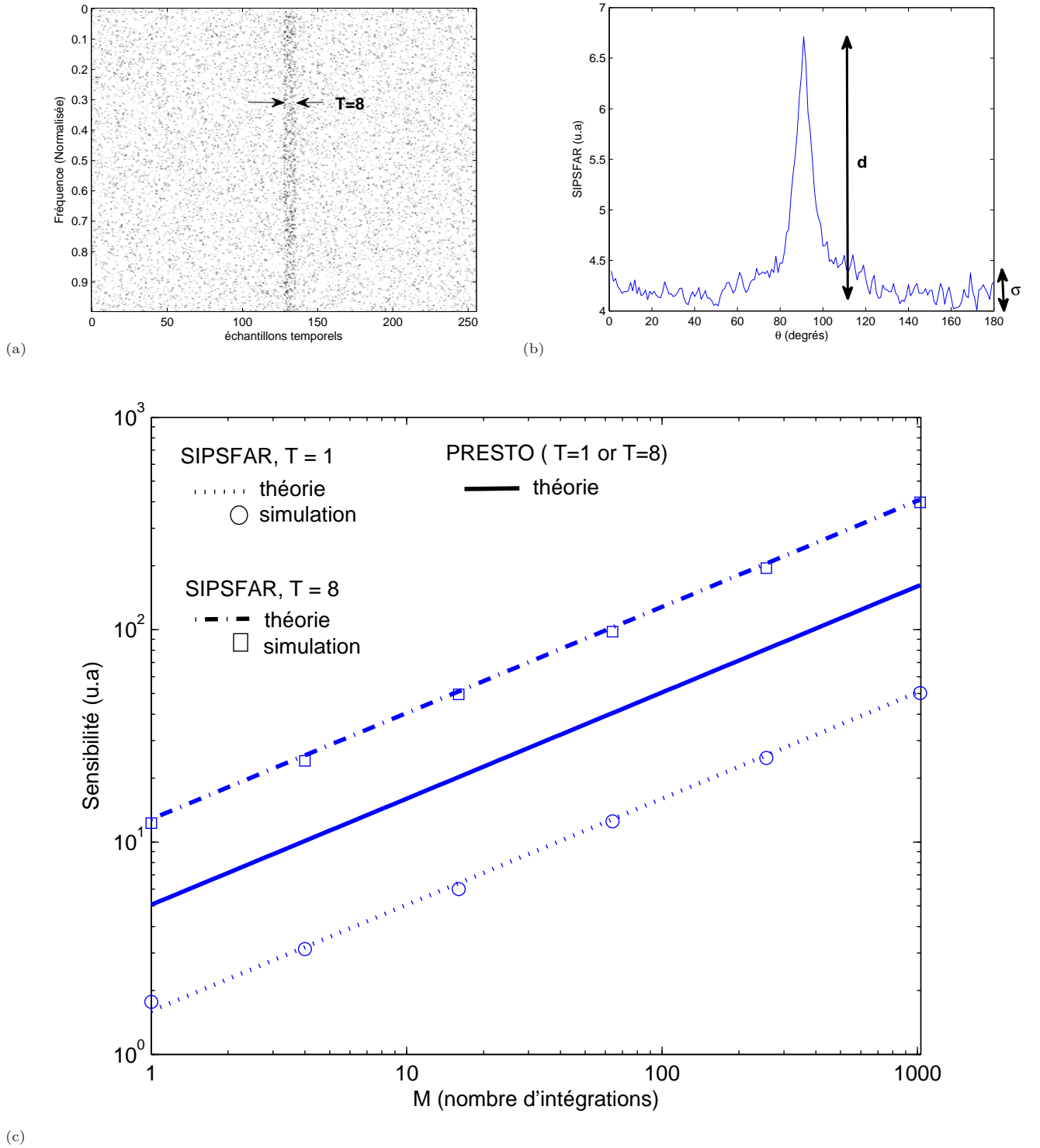


FIGURE 3.6 – Comparaison entre les méthodes SIPSFAR et PRESTO ($N = 256$, $SNR = -5dB$). Deux impulsions de durée $T = 1$ et $T = 8$ échantillons dans le plan t-f sont simulées pour différentes valeurs du nombre d'intégrations M . (a) Exemple d'un plan t-f utilisé pour la simulation. (b) Définition graphique des paramètres utilisés pour le calcul de la sensibilité. (c) Sensibilités pour différentes valeurs de T et de M . Les lignes continues et discontinues correspondent aux valeurs théoriques. Les carrés et les cercles correspondent aux valeurs calculées pour SIPSFAR en simulation.

3.3 Résultats obtenues sur des données du pulsar du Crabe

Les deux détecteurs ont été appliqués sur des données du pulsar du Crabe acquises avec le radiotélescope décimétrique. Cet ensemble de données correspond à 4452 plans T-F générés avec 512 canaux de fréquences. Ces $N = 512$ canaux fréquentiels couvrent la bande de 1.3339 GHz à 1.4619 GHz. La résolution temporelle est égale à $128\mu s$. La durée totale des données est alors 44min 55s. Les paramètres du pulsar utilisés pour la méthode classique sont $DM=56.65 \text{ pc.cm}^{-3}$ et période= 33.63128 ms .

Pour cette observation, on considère :

- Une absence de RFI,
- La période du pulsar est connue,
- PRESTO détecte toutes les impulsions.

La méthode classique est alors utilisée pour identifier les plans T-F avec au moins une impulsion dedans, le critère de détection est fondé sur les deux règles suivantes : "sensibilité supérieure à 4" et "le maximum de l'impulsion est à la phase attendue dans la période du pulsar". Dans la figure 3.7(a), les cercles noirs représentent les 589 plans t-f satisfaisant aux conditions ci-dessus. Cet ensemble de plans T-F est nommé l'ensemble des données *H1*. Les autres 3863 plans T-F ne présentent pas d'impulsion (petits cercles gris) sont nommés l'ensemble des données *H0*. La figure 3.7(a) à droite représente le profil moyen des pulses dé-dispersés à partir des données *H1*.

Dans la figure 3.7(b), les sensibilités obtenues pour les données *H1* entre la méthode classique et SIPSFAR sont comparées. On peut voir que les performances de SIPSFAR sont plus faibles que celles attendues théoriquement (voir 3.2.3). D'une part, ceci peut être expliqué par le fait que les données réelles sont affectées par les effets de propagation dans le milieu interstellaire, tels que la scintillation. D'autre part, le profil réel d'une impulsion n'est évidemment pas rectangulaire comme il était supposé dans la section 3.2.3. Cependant, le point positif est que avec un seuil de détection $S_{SIPSFAR} = 6$, le détecteur SIPSFAR est capable de détecter des impulsions individuelles avec une probabilité de fausse alarme nulle. Plus précisément, SIPSFAR détecte la présence d'une impulsion dans 65 plans t-f avec une sensibilité supérieur ou égale à 6, contre 365 détections avec la méthode classique avec le même seuil. Une première conclusion est que les impulsions très fortes peuvent être détectées sans fausse alarme et sans une connaissance à priori de la mesure de dispersion *DM*. De plus, l'angle correspondant au maximum de détection donne une première estimation du *DM* du pulsar inconnu.

La figure 3.7(b) représente la comparaison entre les sensibilités calculées théoriquement et celles obtenues avec le traitement du pulsar du Crabe. On note que les performances de la méthode SIPSFAR sont systématiquement en dessous des prédictions théoriques. Ceci est expliqué par le fait que dans le calcul théorique les deux hypothèses considérées ne peuvent être complètement vraies dans ce cas. La première est le fait que la puissance de l'impulsion est très inférieure à celle du bruit. Ici, ce n'est pas le cas, car il s'agit des impulsions géantes du Crabe. La deuxième est un peu trop simpliste et ne peut être vraie quel que soit le pulsar observé, l'impulsion est modélisée sous la forme d'un rectangle à la place d'une forme en cloche. Avec cette hypothèse, la sensibilité de SIPSFAR a été artificiellement augmentée d'un facteur *T* (la durée de l'impulsion).

Les figures 3.7(c) et 3.7(d) montrent que la détection avec les seuils les plus faibles peut être atteinte en intégrant les profils individuels. En effet, la figure 3.7(c) montre qu'en intégrant tous les profils individuels correspondants à des données *H1* qui ont une sensibilité plus faible qu'un seuil donné, on obtient un nouveau profil SIPSFAR avec une sensibilité meilleure. Par exemple, en intégrant les 136 premiers profils Radon individuels ayant une sensibilité inférieure à 2.7 (c'est à dire indétectables individuellement), une détection avec zéro fausse alarme via SIPSFAR est obtenue avec une nouvelle sensibilité égale à 6. Dans la figure 3.7(d), on compare un des profils utilisés pour cette intégration (la ligne discontinue en rouge) avec le profil intégré (la ligne continue en bleue). Par conséquent, une deuxième conclusion peut être tirée : SIPSFAR peut donner des résultats positifs de détection des impulsions avec un faible *SNR* en intégrant plusieurs profils individuels exactement comme dans le cas de PRESTO quand on effectue l'empilement des impulsions dé-dispersées. L'avantage, c'est que contrairement à la méthode PRESTO, la connaissance de la période du pulsar n'est pas requise pour SIPSFAR.

L'étude faite sur les données réelles confirme que l'algorithme SIPSFAR est une alternative intéressante pour la détection en temps réel des impulsions individuelles qui sont fortes comme les impulsions géantes et les impulsions individuelles des RRATs.

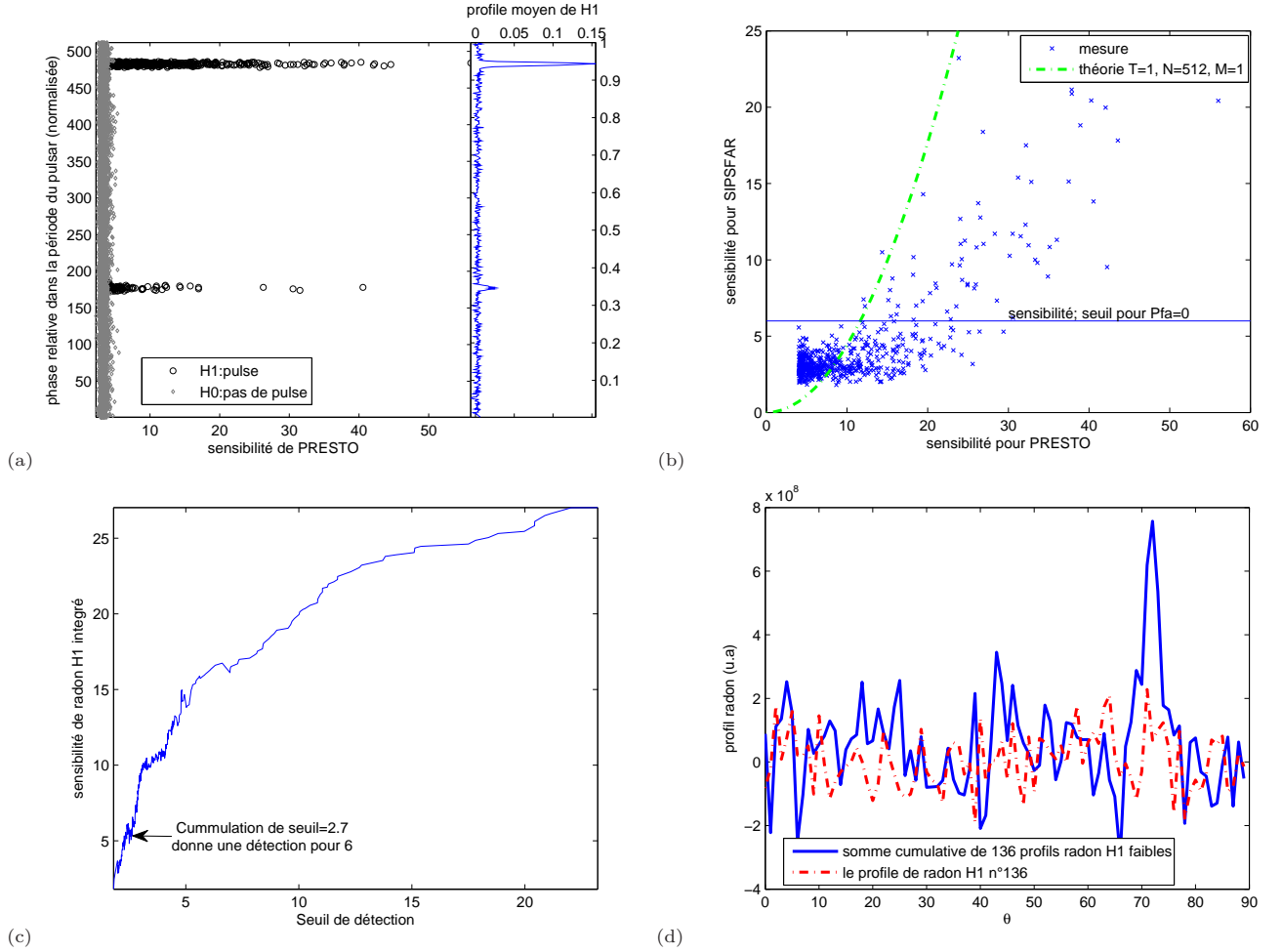


FIGURE 3.7 – Comparaison entre SIPSFAR et PRESTO : (a) Distribution des 4452 plans T-F des données du Crabe. La sensibilité de la méthode classique en fonction de la position du maximum dans la phase du profil du pulsar. Les cercles noirs indiquent la position des détections. Les données des plans T-F correspondant sont libellés $H1$. Les petits cercles gris représentent les données sans pulse dedans, et les données des plans T-F correspondant sont donc libellés $H0$. La figure juste à droite, représente le profil intégré obtenu avec les données $H1$. (b) La comparaison de la sensibilité des deux méthodes testées sur les données $H1$. La ligne discontinue verte représente la relation théorique entre les deux (voir 3.2.3). La ligne horizontale représente la limite inférieure donnant une détection avec une probabilité de fausse alarme (Pfa) égale à zéro. (c) La sensibilité obtenue après intégration de profils individuels ayant une sensibilité supérieure à un seuil donné. (d) Comparaison entre un profil individuel de faible sensibilité et un profil intégré.

3.4 Conclusion

La méthode SIPSFAR est une nouvelle approche pour la recherche des pulsars, combinant à la fois une méthode de recherche efficace en temps réel et certaines capacités d'atténuation des RFI. Elle est basée sur la transformée de Fourier à deux dimensions et la transformée de Radon. Bien que cet algorithme soit moins sensible que la méthode classique basée sur la recherche d'événements dans chaque série temporelle dé-dispersées (de façon incohérente), SIPSFAR peut détecter des impulsions fortes ('single pulses') sans connaître à l'avance le DM du pulsar correspondant.

SIPSFAR pourrait être un outil de dépouillement en temps réel permettant de détecter des événements impulsionnels forts sans a priori. Il pourrait également permettre d'effectuer une première recherche rapide, et probablement en temps réel, dans l'énorme quantité de données acquises lors de la recherche de nouveaux pulsars. C'est une problématique majeure pour la prochaine génération de radiotélescopes.

Chapitre 4

Analyse du survey FOSTER

4.1 Introduction

Dans ce chapitre nous allons appliquer la méthode de détection des événements impulsionnels et intermittents SIPSAR sur un gros volume de données acquises il y a quelques années déjà. L'objectif est de tester cette nouvelle méthode et de la comparer avec la méthode communément utilisée dans l'observation des pulsars. Nous avons montré, dans la comparaison théorique du chapitre 3, qu'elle était plus efficace mais a priori beaucoup plus lourde à mettre en œuvre.

Ce chapitre est composé de trois sections, la première présente les données réelles sur lesquelles le test a été fait. Ensuite, l'implémentation de l'algorithme SIPSFAR va être décrite dans la deuxième section. Les résultats seront présentés et discutés dans la dernière section.

4.2 Le programme de recherche de nouveaux pulsars 'FOSTER'

En 1997–98 avec le radiotélescope de Nançay avant sa rénovation, un programme de recherche de nouveaux pulsars a été mené en collaboration avec R.S.Foster, P.S.Ray (NRL, Washington), D.C.Backer (UC Berkeley) [Foster et al., 1997]. Une instrumentation spécifique a été installée pour cela : le NBPP, Navy Berkeley Pulsar Processor ; sur 96 canaux de 1.5 MHz. Ce programme a été un des premiers à être mené à haute fréquence (1288-1432 MHz) pour limiter les effets de la dispersion. Les données ont été numérisées sur 4 bits avec un échantillonnage à 60 μ s.

Pour chaque observation, les données ont été acquises pendant 2 minutes produisant ~ 2 millions d'échantillons. Sur 2 ans environ, 38971 pointages différents ont été effectués autour du plan galactique vu à Nançay (latitude galactique limitée à ± 5 degrés) sur un total possible de 47229 (Fig. 4.1). Chaque observation de 2 mins a produit un fichier de ~ 100 Mo pour un total de 4.5 To. La première analyse effectuée au NRL à Washington a permis de découvrir deux nouveaux pulsars. Ces pulsars sont maintenant répertoriés comme PSR J0248+6021 ayant une période $P = 217$ ms et un $DM = 376$ pc.cm $^{-3}$, et PSR J2240+5832 ayant une période $P = 139$ ms et un $DM = 264$ pc.cm $^{-3}$ [Foster et al., 1997, Ray et al., 1999, Theureau et al., 2011]. Avec une résolution temporelle de seulement 60 μ s, ces données sont bien adaptées pour tester en grand des recherches systématiques d'événements impulsionnels. A titre d'exemple, la figure 4.2 présente une série d'impulsions individuelles du pulsar J0248+6021 enregistrée lors de l'observation 29298000.

4.3 Recherche d'événements impulsionnels par la méthode SIPSFAR dans un GPU

Des tests de performances ont été réalisés pour prouver l'efficacité des GPUs (Graphic Processing Units) comme une solution appropriée au traitement de données de bande plus large et en temps réel [Desvignes, 2009]. Nous nous proposons donc de porter la méthode de recherche d'événements impulsionnels SIPSFAR sur un GPU. Une carte graphique GTX285 a été installée pour cela sur un PC avec une distribution Linux Gentoo. Ce GPU est doté de 240 calculateurs élémentaires (appelés coeurs), 1 Go de mémoire globale et 16 Mo de mémoire partagée. Pour accéder à la programmation du GPU, la librairie CUDA 2.3 a été installée.

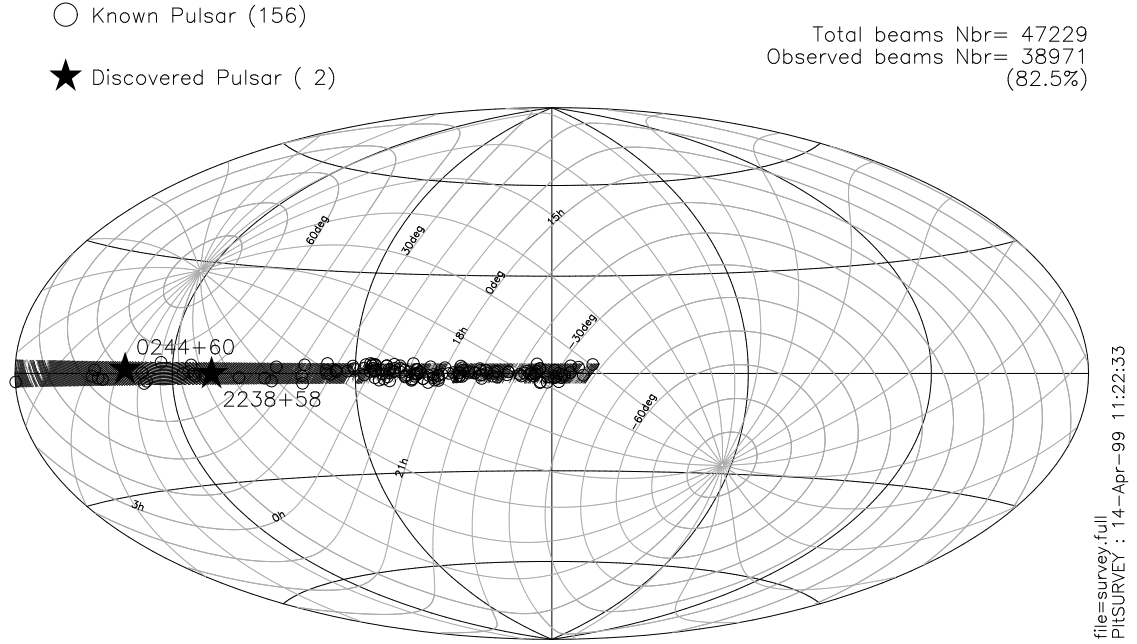


FIGURE 4.1 – Représentation en coordonnées galactiques des 38971 positions observées (sur 47229 potentielles) dans le cadre du programme de recherche de nouveaux pulsars mené en collaboration avec R.S.Foster du NRL Washington à la fin des années 1990. Ce programme a conduit à la découverte de deux pulsars jeunes [Theureau et al., 2011].

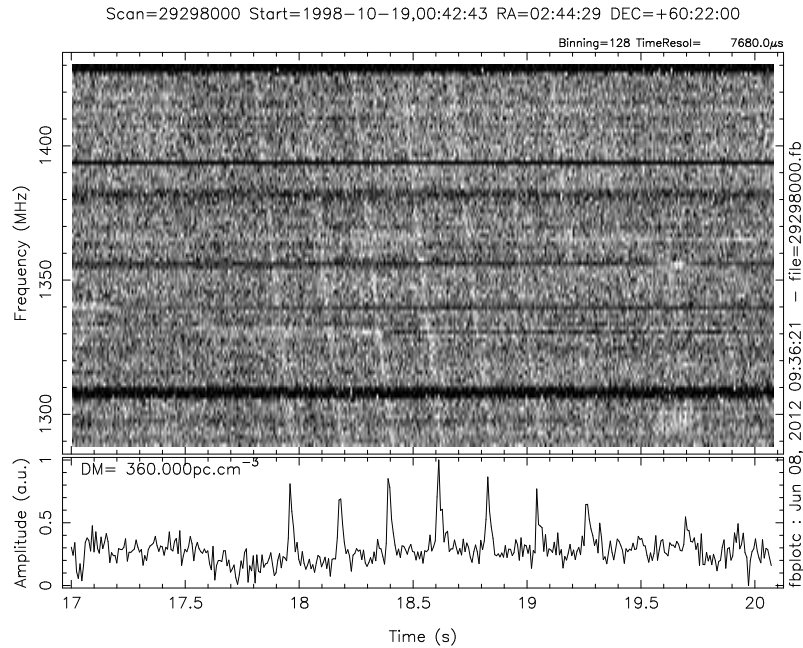


FIGURE 4.2 – Trois secondes de spectre dynamique d’une observation de J0248+6021, un binning sur 128 échantillons a été appliqué en même temps qu’un DM de 360 pc.cm^{-3} pour la série temporelle du bas. Sept impulsions successives sont clairement visibles.

4.3.1 Architecture du GPU et la librairie CUDA

La librairie CUDA exploite au mieux les capacités de calculs parallèles du GPU. Quand le GPU est programmé à travers CUDA C, qui est une extension du langage C, il est vu comme un dispositif de calcul, capable d’effectuer un grand nombre d’opérations parallèles. Une fonction qui doit être exécutée plusieurs fois indépendamment sur différentes données, peut être isolée et compilée comme un programme à part, nommé kernel. Une

fois chargée dans le GPU, il sera exécuté N fois par N "threads" ou tâches parallèlement. Les "threads" peuvent être organisés en une, deux ou trois dimensions, et peuvent être synchronisés à l'intérieur d'un bloc qu'on appelle "thread block". Chaque "thread" est identifié par son ID, qui est son numéro dans le bloc. Finalement, les blocs eux même peuvent être organisés de la même manière dans une grille nommée "Grid" [NVIDIA, 2008]. Le nombre de blocs et le nombre de threads doivent être donnés à l'appel du kernel, tout en respectant certaines règles énoncées dans [NVIDIA, 2008]. La règle la plus importante est que le nombre de registres utilisés par "thread" multiplié par le nombre de threads dans le bloc ne doit pas dépasser le nombre total des registres par multiprocesseur. Sinon, le bloc ne peut pas être exécuté et le kernel correspondant ne va pas se lancer.

La librairie CUDA C englobe le kernel CUFFT qui lance la transformée de Fourier d'un vecteur réel ou complexe, d'une dimension ou plus. Ainsi dans ce paragraphe, on se focalisera sur la mise en œuvre de la transformée de Radon.

En fait, la transformée de Radon qui nous intéresse est simplement la somme des intensités dans l'image suivant des lignes radiales passant par le centre de l'image. Appelons θ l'angle de cette radiale avec l'horizontale. Compte tenu de la symétrie présente dans l'image en sortie de la FFT2D, et pour minimiser le temps de calcul, la transformée de Radon est appliquée uniquement sur le quart de l'image entre 0° et 90° . Alors, pour chaque valeur de θ , l'index des pixels de la ligne radiale correspondante est calculé à l'avance. Ensuite, les index des M angles θ choisis sont stockés dans une matrice qui sera transférée une fois pour toute dans la mémoire partagée du GPU. Cette technique permet d'optimiser les transferts de données et elle améliore la rapidité de traitement. Concrètement, le nombre de blocs est égal au nombre des valeurs de θ choisies, et le nombre de threads est égal au nombre de pixels pour chaque θ (figures 4.3.a et 4.3.b). Ensuite, leur somme est calculée par l'algorithme de réduction expliqué par le schéma de la figure 4.3.c. Le résultat de la somme des pixels le long d'une ligne radiale est récupéré à partir du thread numéro 0 de chaque bloc. Le résultat final est un vecteur de taille M contenant le valeur de l'intensité cumulé du spectre temps-fréquence sur chaque ligne radiale. La détection d'un pic permet une détection en aveugle d'une impulsion dispersée dans le plan temps-fréquence.

La relation reliant l'angle de Radon à la mesure de dispersion n'est pas linéaire, l'angle de Radon dépend en effet du DM, de la fréquence centrale d'observation f_0 , de la bande passante totale BW , de l'échantillonnage temporel dt ainsi que d'une fonction trigonométrique. L'angle θ est donné à partir de l'équation 3.10 :

$$\theta = \arctan\left(DM \frac{8.6 \times 10^6}{N_{chan} \times n_{bin} \times dt} \frac{BW}{f_{ctr}^3}\right)$$

Il est aisé de voir qu'un angle similaire peut être obtenu pour diverses combinaisons de DM et dt . Ainsi pour accéder à des mesures de dispersion DM élevées, il est nécessaire d'appliquer des facteurs de 'binning' croissants pour augmenter dt . Les échantillons temporels successifs sont alors moyennés sur 8, 16, 32, 64 ou 128 valeurs. Ceci permet de couvrir plus efficacement l'ensemble des valeurs de DM comme le démontre la figure 4.4. Si tous les facteurs de binning sont utilisés, les limites 'théoriques' des angles de Radon à utiliser pour détecter de façon unique un événement sont $[45^\circ - 63^\circ]$. La limite 45° est choisie pour avoir le maximum de la puissance de l'impulsion qui est obtenue pour une inclinaison maximale de 45° (toute l'impulsion est dans le plan t-f). La limite 63° est obtenue là où le dernier DM a été calculé avec le n_{bin} précédant afin d'éviter des détections multiples.

Nous verrons plus tard que ceci n'est pas facilement applicable sans inconvénient important.

4.3.2 Application aux données dites 'FOSTER'

Une suite de codes a été écrite pour analyser le gros volume que constitue les données de recherche de pulsars dites 'FOSTER' de façon automatique, nous en ferons une description générale du plus bas niveau vers le plus haut.

Un premier programme applique des transformées de Radon successives sur un fichier de données pour un facteur de 'binning' temporel donné et fournissant le profil Radon (un vecteur de $M = 128$ valeurs pour différents angles), la somme de tous les profils Radon est également sortie. Lorsque pour un certain angle l'intégrale dépasse un niveau de référence, le programme sort le profil Radon intégré dans un fichier nommé de façon unique et adéquate. C'est un second programme plus général qui lance le premier programme pour différentes valeurs de 'binning' sur l'ensemble des fichiers de données listés dans un fichier, ce fichier sera modifié pour y indiquer la date du traitement et éviter de le refaire en cas d'arrêt de ce traitement.

Ce second programme lance également un programme en charge de produire de petites images des transformées de Radon ainsi que de stocker tous les événements impulsionnels dans un fichier unique qu'il sera facile d'utiliser pour les comparaisons de la suite.

Le traitement de l'ensemble des données (~ 40000 observations) a été effectué sur le PC équipé d'un GPU en un tout petit peu plus d'un mois. Ceci sans que le code ne soit particulièrement optimisé. Par exemple, le fichier de données est relu depuis le disque dur systématiquement pour les cinq différentes valeurs du facteur

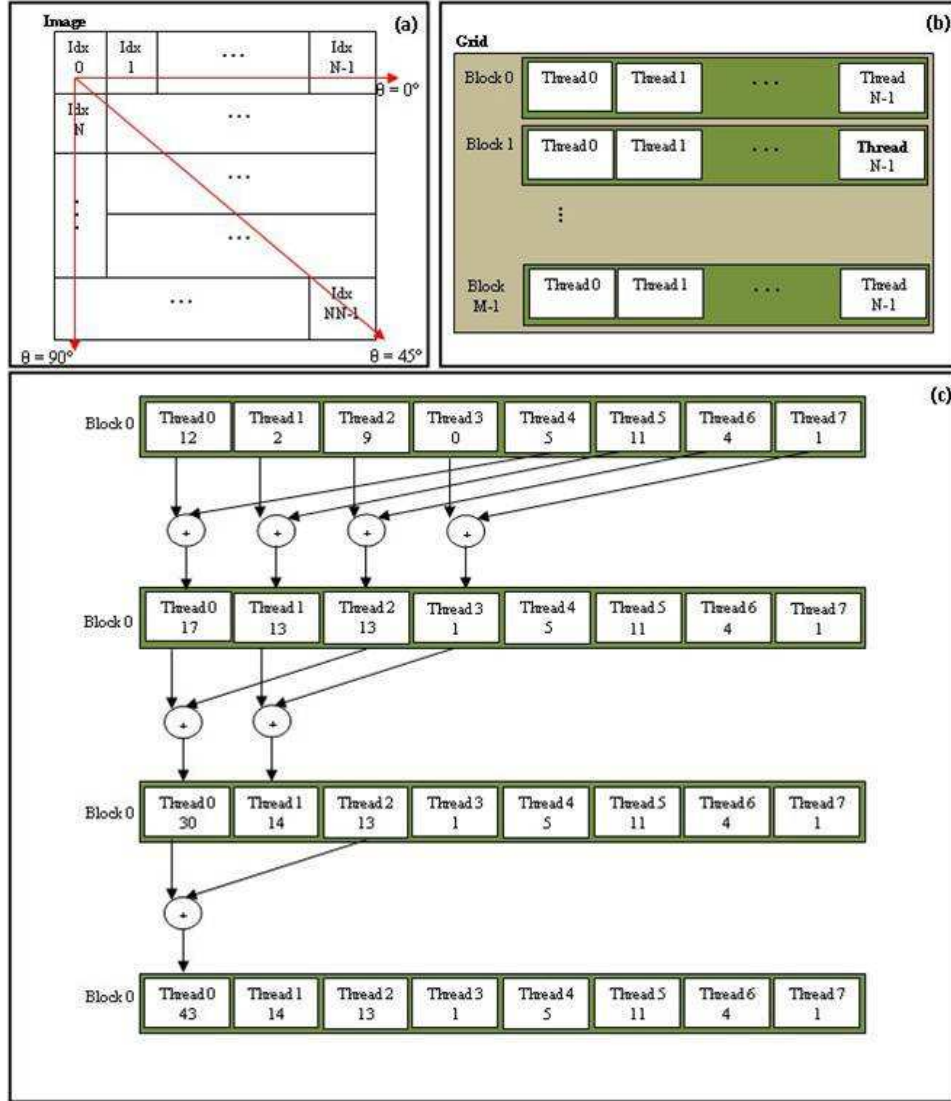


FIGURE 4.3 – Architecture CUDA de l'exécution du kernel 'RADON'

de binnig. Ne faisant que ~ 100 Mo compressé, il pouvait être envisagé de stocker son contenu en mémoire et gagner ainsi du temps (à ce niveau de performance des GPUs, la mise en mémoire des données n'est plus une fraction négligeable du temps d'exécution). Le traitement PRESTO standard (qui cherche aussi les signaux périodiques dans le domaine de Fourier, mais dont la majeure partie du travail est de produire les séries temporelles dédispersées utilisées pour la recherche d'événements impulsionnels) a nécessité environ 3 mois de calcul sur un cluster de 64 nœuds. C'est donc probablement par un facteur 200 à 500 que la méthode SPISFAR est plus rapide par rapport à la méthode PRESTO standard.

4.4 Comparaisons des résultats SPISFAR et PRESTO

Avant d'envisager de traiter de façon statistique les événements détectés par les deux méthodes et de les comparer, il a fallu en normaliser les résultats.

Les événements PRESTO sont d'abord nettoyés en ne gardant que les observations ayant produit moins de 5000 événements de façon à éliminer les observations remplies de parasites (une des observations corrompues par des parasites sera traitée particulièrement plus bas).

Du fait de sa largeur temporelle non négligeable et de la proximité des nombreux DMs utilisés (environ 1500) lors de la recherche de pulsars avec PRESTO, un même événement impulsionnel apparaît en général pour de nombreux DMs voisins. Il est nécessaire de parvenir à (ou de s'approcher de) ne garder qu'un seul

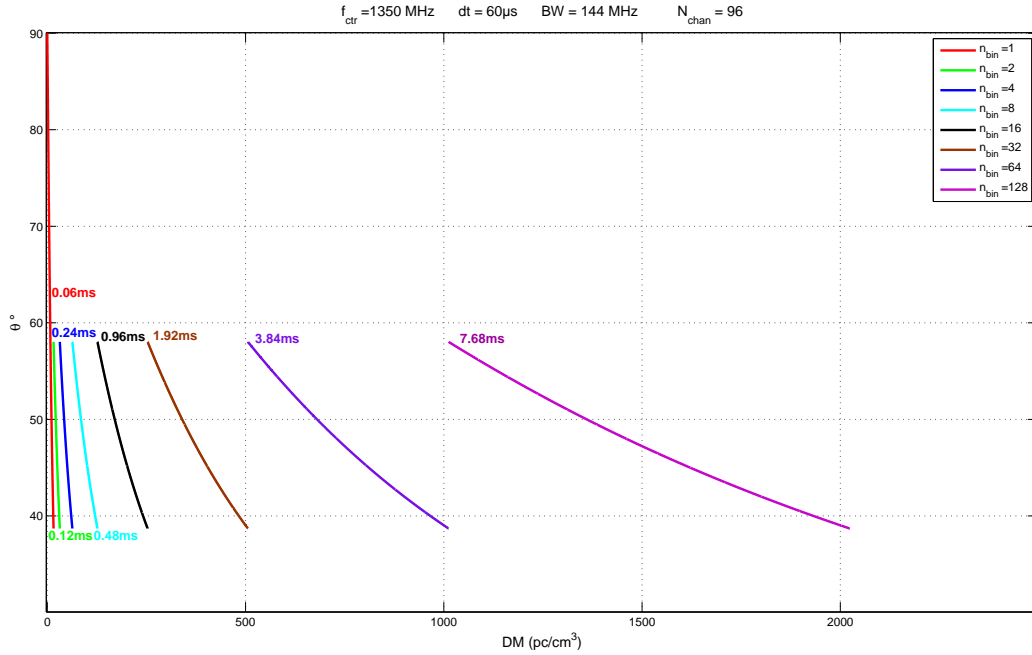


FIGURE 4.4 – Evolution de l'angle de Radon en fonction de la mesure de dispersion DM et du facteur de binning imposé. Pour chaque facteur de binning, les courbes ont été limitées pour éviter le recouvrement et donnent les angles de Radon à potentiellement utiliser pour éviter les détections multiples.

événement pour faciliter les comparaisons suivantes. Un algorithme simple permet de regrouper les différents événements détectés à différents DMs et relevant manifestement du même événement impulsif. Il s'agit de passer en revue tous les événements en les regroupant au fur et à mesure dès que la distance à un groupe en cours de formation (et par rapport à son barycentre) est inférieure à un seuil donné (30 pc.cm^3 en DM et 0.1 sec en temps). Un exemple de l'efficacité de ce regroupement sera visible lors de l'examen de l'observation 29298000 effectuée sur le pulsar J0248+6021.

Malgré l'élimination, dès le calcul de la transformée de Radon, des axes correspondants aux angles 0° et 90° (cf. 3.2 et figure 3.3), il a été noté des détections parasites à très petit ou très grand angle. Les événements SIPSFAR ont donc été nettoyés en ne gardant que les événements correspondants à des angles compris dans l'intervalle $[2^\circ - 87^\circ]$. De cette façon, le nombre total d'événements SIPSFAR est passé de 1115576 à 548394 (la moitié des événements environ a été éliminée). En n'appliquant pas la règle stricte $[45^\circ - 63^\circ]$, nous nous gardons ici la possibilité de détecter le même événement avec des facteurs de binning différents.

Nous commencerons par regarder trois cas particuliers, une observation très bruitée et deux observations de pulsars connus pour émettre des impulsions individuelles, avant d'aborder une étude statistique simple.

4.4.1 Une observation très parasitée

L'observation 13198010 est particulièrement bruitée, comme le montre la figure 4.5 qui présente environ 1/4 de seconde du spectre dynamique de cette observation. Certains canaux sont complètement saturés et de forts parasites (qui saturent le convertisseur en apparaissant en dessous de la moyenne) apparaissent régulièrement dans les données. Basée sur l'analyse dans le domaine temporel, la méthode PRESTO produit évidemment un nombre 'astronomique' de détections, avec plus de vingt millions d'événements pour cette seule observation. La figure 4.6 produite par le code PRESTO montre un espace temps-DM complètement saturé d'événements. Des paliers sont visibles aux endroits des changements de 'binning' opérés à des valeurs particulières de DMs pour accélérer les calculs (ces changements modifient évidemment les conditions de détection des événements et donc la 'densité' d'événements). Basée sur une analyse dans le domaine fréquentiel, la méthode SIPSFAR ne produit que deux événements à petit DM (vers 50 et 115 secondes à partir du début de l'observation, Fig. 4.7). Même sans aucun événement détecté (et c'est impossible ici !), ce résultat montre que la méthode SIPSFAR est davantage robuste aux parasites et ne gaspillera pas inutilement trop de temps de calcul.

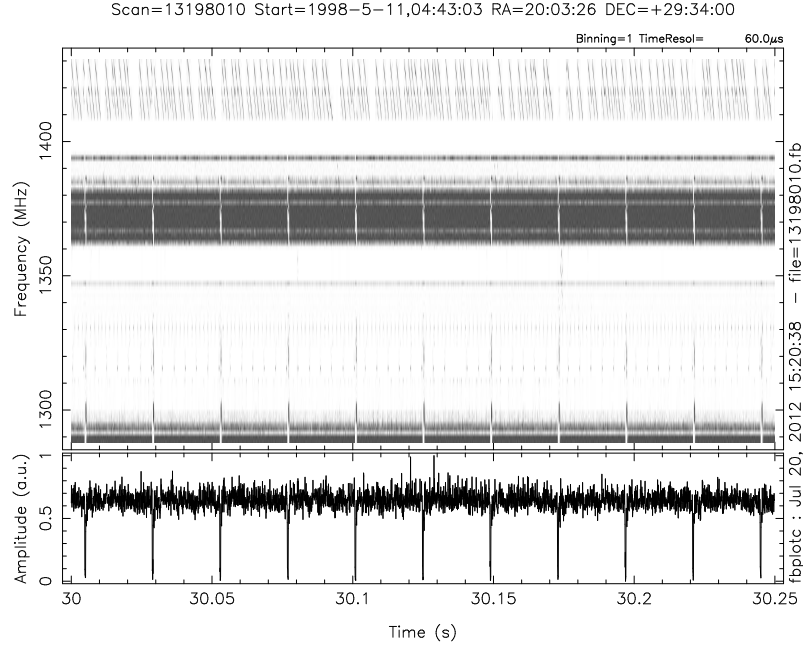


FIGURE 4.5 – Exemple de spectre dynamique obtenu sur une portion de l’observation emplie de parasites.

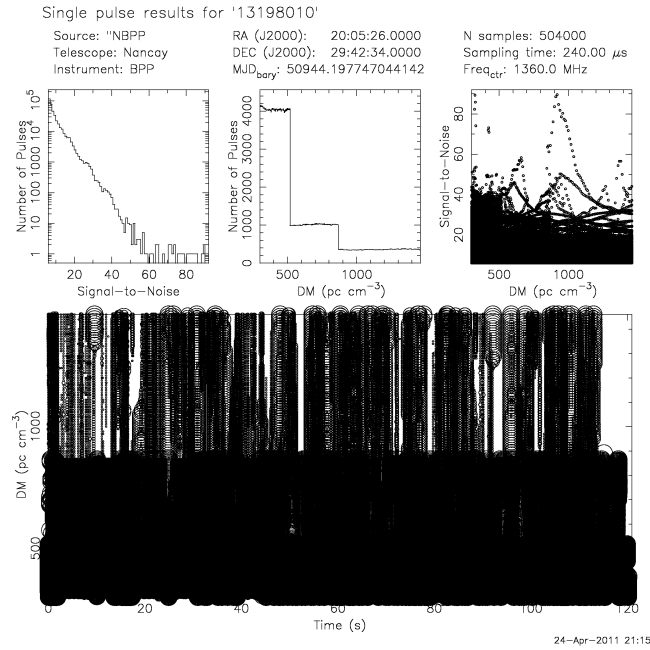


FIGURE 4.6 – Détections d’événements impulsionnels par PRESTO montrant une saturation irrécupérable. La technique PRESTO montre là une limite importante avec 20680710 événements.

4.4.2 Le pulsar du Crabe B0531+21

Le pulsar du Crabe est connu pour émettre des impulsions intenses très étroites, et ayant un $DM = 56.7 \text{ pc.cm}^{-3}$. c’est même la façon dont il a été découvert à la fin des années 1960 ([Staelin and Reifenstein, 1968]). Le pulsar du Crabe a été observé une fois pendant le programme de recherche FOSTER (numéro 21198027). Dans l’observation correspondante, la méthode SIPSFAR ne détecte aucun événement au bon DM quand PRESTO en voit trois (voir figure 4.8).

Il y a plusieurs explications possible pour expliquer cette non-détection SIPSFAR et le peu de détection PRESTO. Ces observations ont été effectuées avec le radiotélescope avant sa rénovation (5 fois moins rapide que maintenant avec une sensibilité 2.2 fois plus faible) et avec une instrumentation d’une bande totale effective

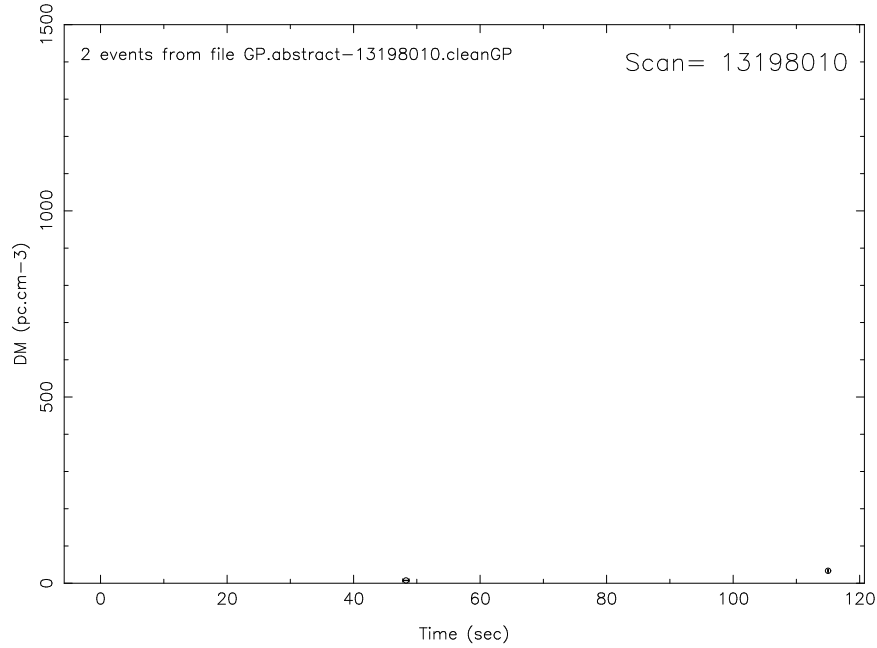


FIGURE 4.7 – Détections d’événements impulsionnels par la méthode RADON dans l’observation très parasitée. Ici, seuls deux événements (un à presque 50 sec et l’autre à 115 sec pour un DM proche de 0) sont détectés et démontre que la méthode est relativement insensible aux parasites.

de ~ 100 MHz (plusieurs canaux étaient inopérants), or l’on sait que la méthode SIPSFAR est moins efficace lorsque le signal est faible. Contrairement aux impulsions du pulsar PSR J0248+6021, les impulsions du Crabe sont extrêmement étroites (quelques microsecondes vs. quelques millisecondes), et auront donc davantage de difficulté à ressortir avec une résolution de $60\mu\text{s}$ et encore plus après avoir effectué des ‘binning’ importants. Toutes ces raisons expliquent probablement pourquoi la méthode SIPSFAR échoue ici, d’autant que nous n’avons qu’une observation pour effectuer ce test et ne pouvons donc pas espérer être aidé par la scintillation qui parfois amplifie le signal radio de façon importante (et l’atténue aussi mais ce n’est pas ce qui est espéré ici).

4.4.3 Le pulsar PSR J0248+6021

Le pulsar J0248+6021 découvert lors de cette recherche de la fin des années 1990 s’avère être un émetteur d’impulsions individuelles prolifique. Ceci est bien visible lors de l’observation correspondant au scan numéro 29298000 présentée en début de chapitre (Fig. 4.2). La figure 4.9 montre les détections SIPSFAR et PRESTO sur la même intervalle de temps. En comparant les parties haute et basse de la figure, nous noterons que l’algorithme de regroupement des événements PRESTO fonctionne de façon satisfaisante. Deux transformées de Radon sont tracées en figure 4.10 avec un exemple de détections multiples (quatre événements dans une fenêtre large) avec un grand SNR et un exemple d’une détection unique (un seul événement dans une fenêtre étroite) avec un petit SNR. En effet, du fait des différents facteurs de ‘binning’ utilisés avant la FFT 2D et le calcul de la transformée de Radon, la méthode SIPSFAR détecte des impulsions soit individuellement (lorsque le binning est faible et l’impulsion forte), soit en groupe (lorsque le binning est grand et que plusieurs événements relativement faibles se suivent et se combinent). Le dernier événement SIPSFAR (vers ~ 19 sec) montre bien que lorsque plusieurs événements faibles se suivent et entrent dans la fenêtre temporelle de la méthode, il peut y avoir détection. Ceci alors que seul le premier de ces événements (le plus fort) est détecté avec un binning plus petit. Si nous avions restreint la détection aux angles de Radon $[45^\circ - 63^\circ]$, nous n’aurions pas forcément gardé les détections aux petits angles pour de grands facteurs de binning qui recouvrent de nombreux événements faibles. Même si cela ne facilite pas la comparaison ultérieure avec la méthode standard de PRESTO, ceci montre bien qu’il y a des raisons importantes pour prendre des distances avec cette règle un peu trop théorique.

Les événements respectivement détectés par les méthodes SIPSFAR et PRESTO dans cette observation sont maintenant analysés plus en détail. La Table 4.1 donne les événements SIPSFAR dans l’ordre chronologique avec l’indication explicite de la largeur de la FFT utilisée et qui doit être vue comme la fenêtre dans laquelle le

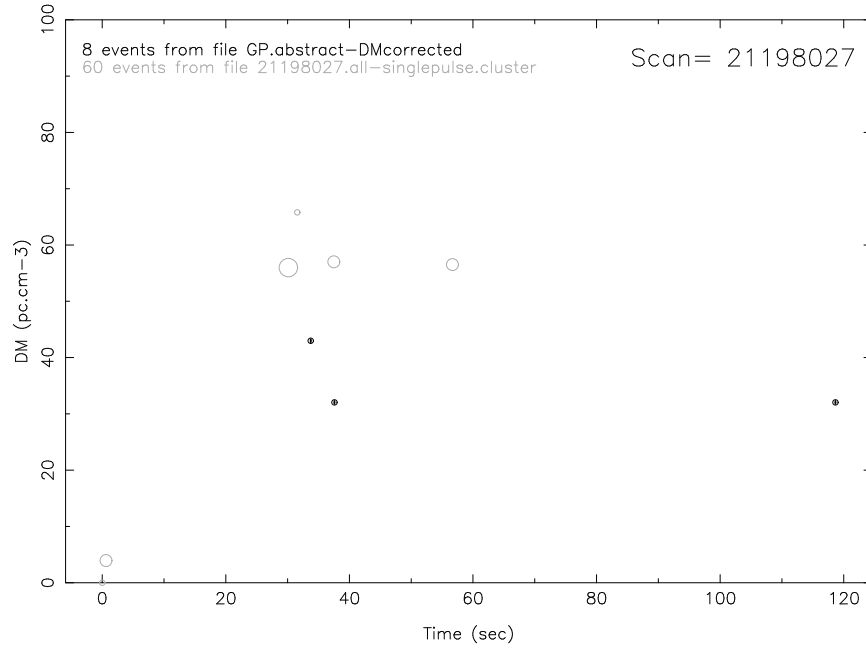


FIGURE 4.8 – Superposition des impulsions détectées par PRESTO (en gris) et par RADON (en noir) dans une observation du pulsar du Crabe (B0531+21, scan=21198027). Les trois impulsions assez faibles détectées par PRESTO ne le sont pas par SIPSFAR. A la date des observations, le radiotélescope n’avait pas encore été rénové et présentait une sensibilité 2.2 fois plus faible expliquant probablement la non-détection par la méthode SIPSFAR, car celle-ci a de faibles performances sur les signaux les plus faibles.

(ou les) événement(s) ont été détectés. La Table 4.2 donne elle les événements PRESTO, toujours dans l’ordre chronologique, et mentionnant, en dernière colonne, le nombre d’événements qui ont été regroupés lors de la normalisation.

La Table 4.3 récapitule le nombre d’événements détectés par chacune des deux méthodes à comparer : PRESTO et SIPSFAR. Une recherche automatique, détaillée dans la prochaine section, des plus proches distances entre d’événements PRESTO et SIPSFAR a été menée. Elle a permis de trouver que les 9 événements apparaissant à la bonne mesure de dispersion DM (entre ~ 340 et ~ 420 pc.cm^{-3} environ) trouvés par SIPSFAR ont également été détectés par la technique PRESTO. A l’inverse, parmi les 30 événements PRESTO, il a été trouvé 7 associations avec des événements SIPSFAR uniques avec jusqu’à 10 événements PRESTO associés sans se préoccuper du fait qu’un événement SIPSFAR puisse être mentionné de multiple fois. Avec 7 événements SIPSFAR retrouvés de façon unique dans les données PRESTO, il se trouve 2 événements, les SIPSFAR n° 4 et 5, sans associations PRESTO. Il sera facile de voir que ces 2 événements sont à petit facteur de binning (16 et 32 respectivement) et sont en fait complètement recouverts par d’autres détections SIPSFAR (c’est bien visible en haut à gauche de la figure 4.9).

Notons un événement PRESTO apparaissant à un DM largement en dessous de 300 pc.cm^{-3} à 17.5 secondes bien visible sur la figure 4.9. Du fait de son faible DM, il n’a pas été retenu dans l’analyse et n’est de toutes façons pas visible dans la figure 4.2.

Si on ne se restreint pas aux angles $[45^\circ - 63^\circ]$ (cf. 4.3.1), il ne reste alors plus que les événements n° 5 et 6 de la Table 4.1 (éventuellement les n° 2 et 6 si on relaxe de quelques degrés). Ceci n’est manifestement pas satisfaisant car beaucoup d’événements sont alors manqués.

L’analyse de ce cas particulier révèle bien que la comparaison entre les méthodes SIPSFAR et PRESTO n’est pas aisée. Les critères définissant un événement sont différents, le point le plus délicat étant la fenêtre temporelle de SIPSFAR dans laquelle peuvent se glisser de nombreux événements PRESTO.

L’observation scannée 15798046 (Fig. 4.11) permet de confirmer que la méthode SIPSFAR détecte de façon satisfaisante les événements relativement intenses tout en étant moins sensible que la méthode mise en œuvre dans PRESTO.

	SNR	DM	Time interval	binning	angle de Radon
1	7.300	364.046	11.704320-12.679680	128	77.3
2	6.331	359.308	17.556480-18.044161	64	66.1
3	12.233	364.046	17.556480-18.531839	128	77.3
4	5.849	402.589	17.678400-17.800320	16	26.7
5	5.508	405.293	18.044161-18.288000	32	45.0
6	5.817	383.379	18.044161-18.531839	64	64.7
7	6.559	379.125	18.531839-18.653761	16	28.1
8	9.651	406.082	18.531839-19.507200	128	75.9
9	6.152	385.871	23.652479-23.896320	32	46.4

TABLE 4.1 – Liste des 9 événements SIPSFAR de l’observation 29298000 correspondants à des impulsions apparaissant à des DMs voisins de celui de J0248+6021 sur les 50 détectés au total.

	SNR	DM	Time groupe	Taille		SNR	DM	Time	Taille
1	9.34	358.00	11.940578	52	16	9.30	371.56	61.229424	144
2	7.38	360.69	12.160028	48	17	6.25	358.00	69.253685	3
3	9.03	347.28	17.807737	61	18	7.16	366.43	69.464775	21
4	7.14	357.35	18.021168	37	19	7.46	392.50	70.545784	26
5	6.69	358.60	18.884592	15	20	6.04	406.00	72.480721	1
6	10.02	380.63	18.229382	46	21	9.95	370.55	82.841415	216
7	10.42	389.90	18.445192	106	22	7.77	395.34	89.869034	47
8	8.34	400.57	18.660944	58	23	7.45	375.66	93.780373	35
9	6.49	373.23	19.105883	13	24	7.39	368.57	94.650475	23
10	8.90	379.19	23.872293	31	25	7.15	414.23	101.317055	39
11	8.01	342.87	31.923248	47	26	7.08	380.17	103.330589	18
12	6.02	370.00	53.401680	1	27	9.35	391.51	103.545166	84
13	6.90	380.14	54.051910	21	28	6.12	367.00	111.803574	4
14	6.82	362.09	57.743481	11	29	6.39	347.50	113.585876	2
15	8.45	387.35	57.954426	46	30	6.79	356.31	117.228928	16

TABLE 4.2 – Liste des 30 événements PRESTO correspondants à des impulsions réelles apparaissant à des DMs entre 340 et 420 pc.cm^{-3} sur les 149 détectés au total. Notons que les 9 événements n° 2 à 10 correspondent aux 9 événements SIPSFAR listés en Table 4.1

Nombre total d’événements PRESTO	149
Nombre d’événements pulsar vus par PRESTO	30
Nombre d’associations uniques avec SPISFAR	7
Nombre d’associations avec SPISFAR	9
Nombre de SIPSFAR manquants	2
Nombre d’événements SIPSFAR	50
Nombre d’événements pulsar vus par SIPSFAR	9
Nombre d’associations avec PRESTO	9

TABLE 4.3 – Associations d’événements PRESTO - SIPSFAR pour l’observation 2929800 de PSR J0248+6021. Comme attendu à partir des sensibilités théoriques des deux méthodes, tous les événements SIPSFAR sont associés à des événements PRESTO alors que seuls 9 événements PRESTO sont associés à des événements SIPSFAR.

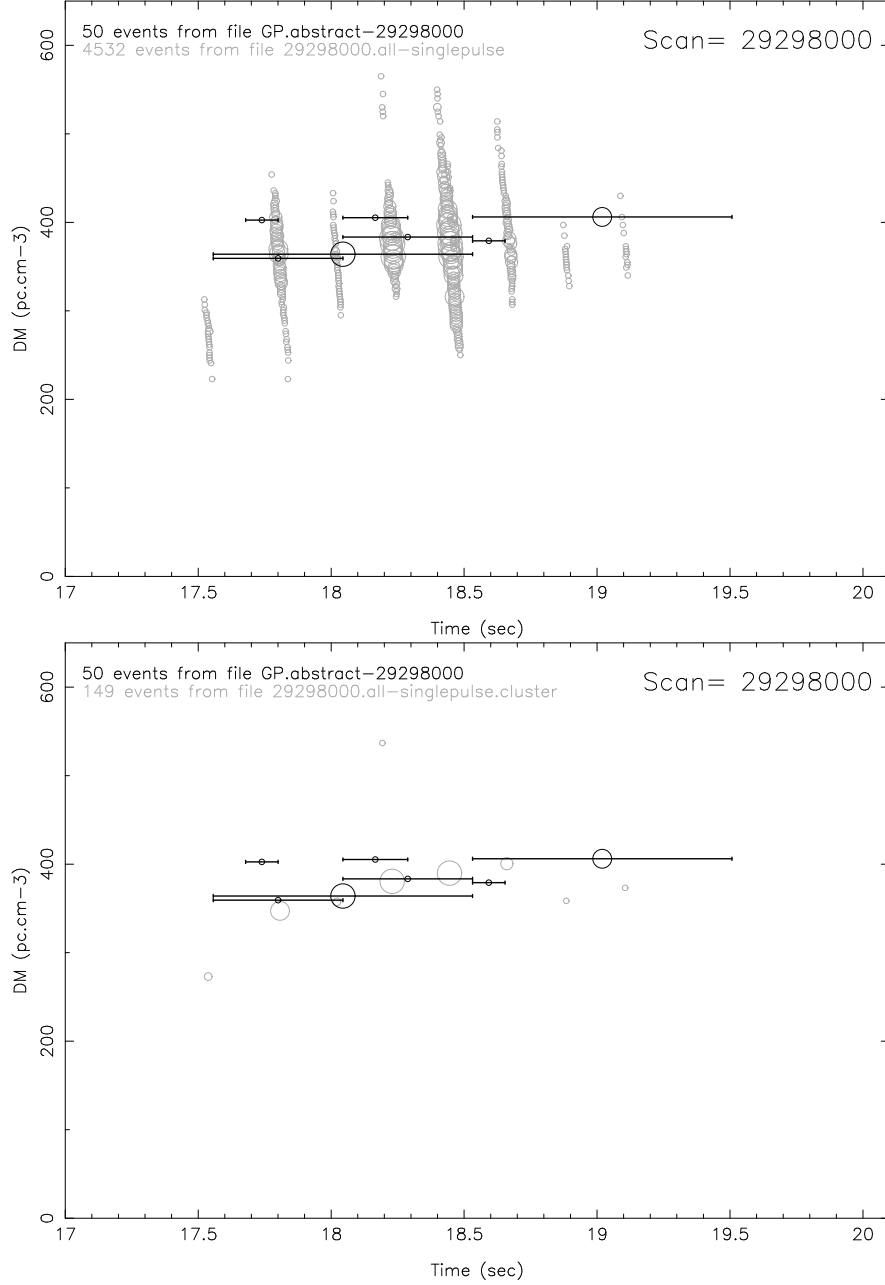


FIGURE 4.9 – Superposition des impulsions dispersées détectées par PRESTO (en gris avec des détections à différents DMs) et par SIPSFAR (en noir avec une barre horizontale indiquant la largeur de la fenêtre de la FFT). Trois secondes d’une observation du pulsar J0248+6021 (scan 29298000) montrent sept (ou huit ?) impulsions successives toutes détectées également par RADON sauf la première alors que les trois dernières sont détectées toutes ensemble. (haut) tous les événements PRESTO sont représentés (bas) un regroupement des événements PRESTO a été effectué. On pourra utilement comparer avec la figure 4.2 qui montre la même portion de données.

4.4.4 Un traitement statistique simple

Malgré les limitations pointées dans la section précédente à la comparaison détaillée des deux méthodes SIPSFAR et PRESTO, nous allons tenter une analyse statistique simple.

Après la première phase de ‘nettoyage’ et l’étude de quelques cas, nous avons donc systématiquement cherché les plus courtes distances en partant d’un événement SIPSFAR. Pour commencer, pour une observation donnée et pour chaque événement SIPSFAR, la distance avec tous les événements PRESTO est calculée ($\sqrt{dDM * dDM + 500 * dt * dt}$) avec dDM l’écart en mesure de dispersion DM et dt l’écart en temps d’appar-

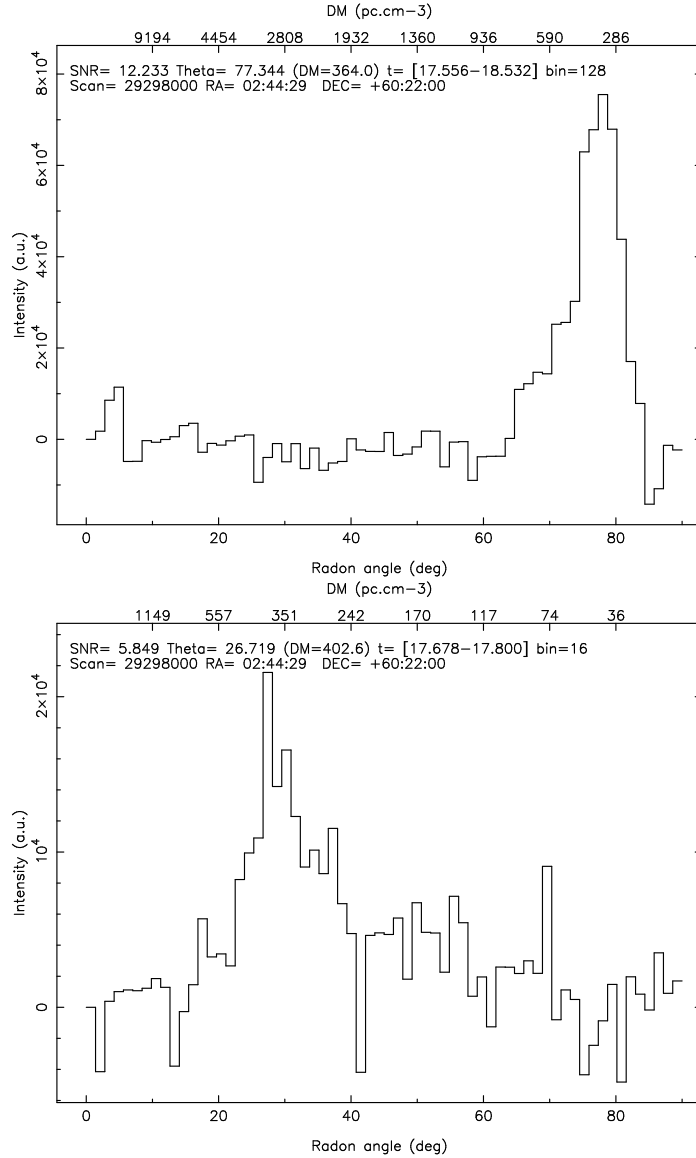


FIGURE 4.10 – Profils Radon de deux événements impulsionnels détectés sur le pulsar J0248+6021 (scan=29298000) pour différent facteur de 'binning'. (haut) quatre événements détectés simultanément (n° 2, 3, 4 et 5 de la figure 4.2) avec une fenêtre large. (bas) un événement isolé (n° 2) détecté avec une fenêtre étroite.

rition de l'événement (le milieu de la fenêtre pour SIPSFAR). Pour cet événement SIPSFAR, seule la valeur de distance la plus faible est conservée ainsi que le numéro de l'événement PRESTO. Dans le calcul de la distance, le coefficient 500 sur le temps permet d'équilibrer le poids des deux dimensions présentes dans cette distance.

A la fin de ce traitement, nous avons un ensemble de couples (dDM,dt) liants les événements SIPSFAR et PRESTO les plus proches. C'est à partir de ces données que seront effectués divers traitements maintenant détaillés.

Pour commencer, la façon la plus naturelle d'obtenir une vue rapide du résultat est de placer un point pour chaque couple obtenus dans un plan (dt,dDM) et de chercher une concentration autour du point central de distance nulle (Fig. 4.12). Notons que même en l'absence d'événements réels détectés par les deux méthodes, le fait qu'il existe en général plusieurs événements SIPSFAR et PRESTO par observation, produit une concentration des distances autour de zéro. En effet, pour un événement SIPSFAR donné, il sera toujours possible de trouver un événement PRESTO pas trop loin, ni en DM, ni en temps. Selon l'écartement moyen (en DM et en temps) des événements PRESTO, nous obtenons donc un fond maximum en (0,0) et au-dessus duquel se superpose une distribution plus piquée correspondant aux événements réels liés à une même impulsion détectée par les deux méthodes. Toute la question est d'essayer de séparer au mieux les deux distributions.

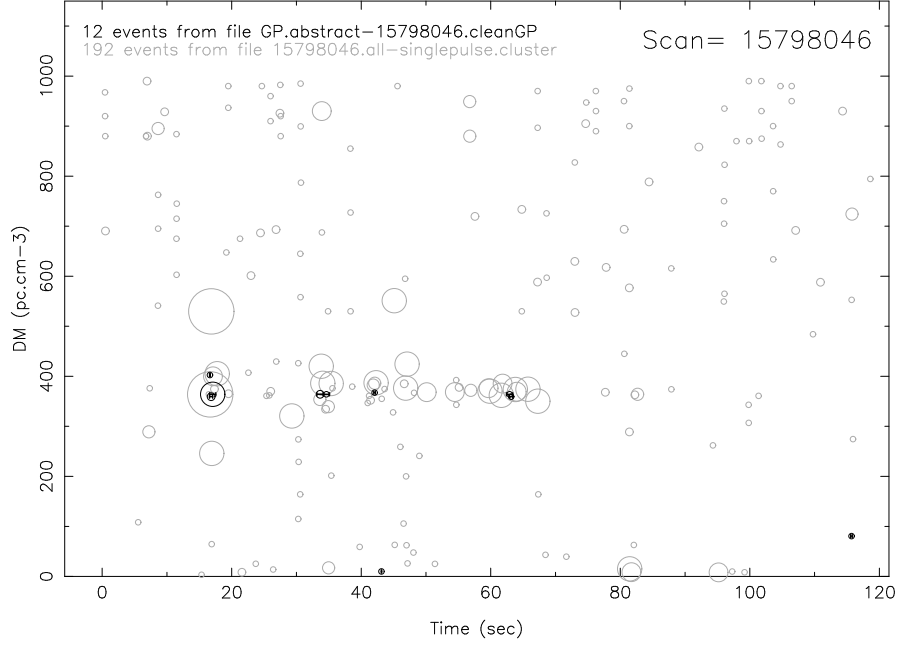


FIGURE 4.11 – Superposition des impulsions détectées par PRESTO (en gris) et par SIPSFAR (en noir) dans l’observation de PSR J0248+6021 (scan 15798046). Si SIPSFAR détecte moins d’évènements comme prévu car la technique est moins sensible, elle semble également détecter moins d’évènements parasites.

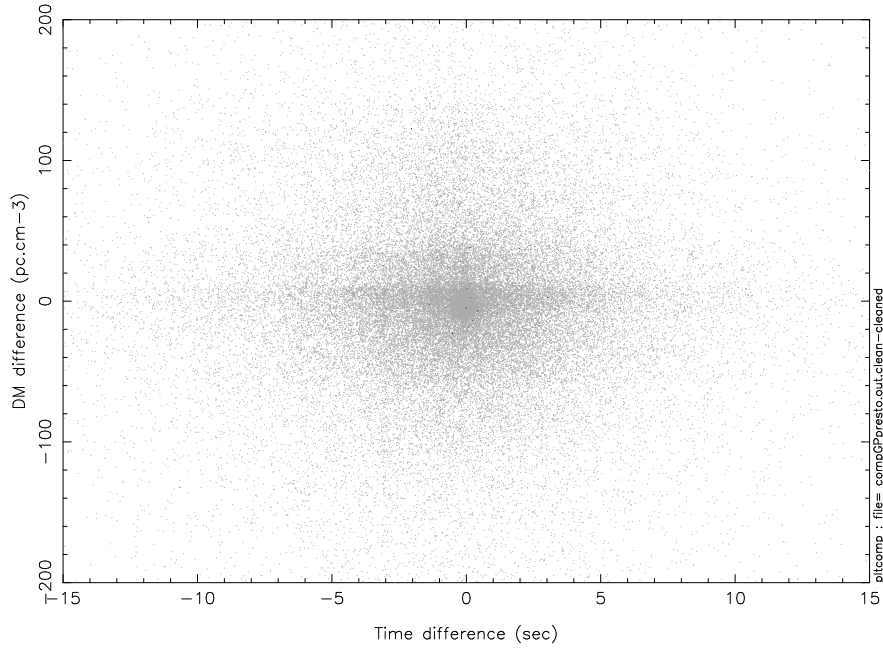


FIGURE 4.12 – Visualisation des différences de DM et de temps pour les couples d’évènements (SIPSFAR,PRESTO) les plus proches.

Pour explorer plus en détail la structure de la distribution des différences (DM, t) , il est intéressant de tracer l’histogramme selon les différences de DM sans ou avec une limitation sur les différences en temps (Fig. 4.13). Comme attendu les histogrammes $dn\ dDM$ obtenus avec une sélection sur les différences en temps (ne sont gardées que les différences inférieures à 0.125 et 0.5 secondes en valeur absolue) sont plus piqués et présentent moins la distribution plus large provenant des associations par hasard.

Dans l’autre direction, ce sont des histogrammes en différence de temps qui sont représentés en figure 4.14. Bien que de façon moins prononcée, les histogrammes en différence de temps obtenus avec des filtrages sur les différence de $|DM|$ limitées à 10 et 50 pc.cm^{-3} en valeur absolue présente moins d’évènements et une diminution

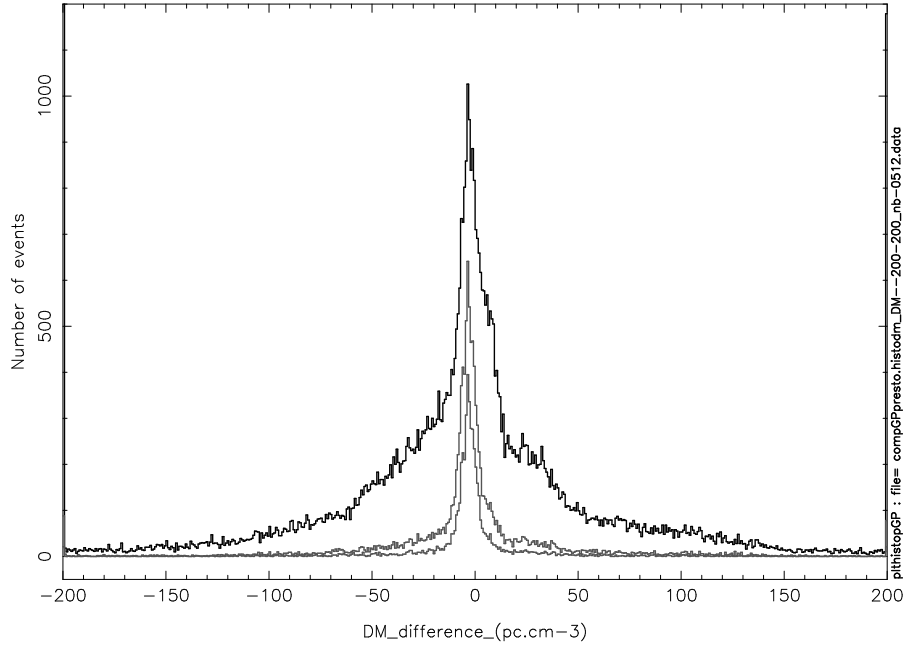


FIGURE 4.13 – Histogrammes des différences en DM : (noir) toutes les différences en temps sont prises en compte, (gris) les différences temporelles inférieures à 0.125 et 0.5 secondes.

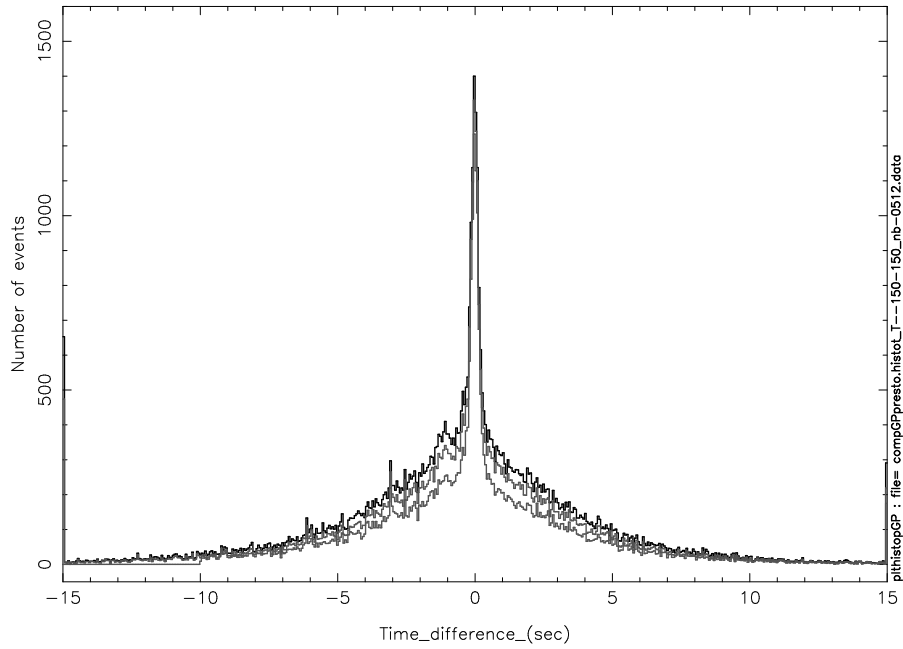


FIGURE 4.14 – Histogrammes des différences en temps : (noir) toutes les différences en DM sont prises en compte, (gris) les différences en DM inférieures à 10 et 50 pc.cm^{-3} .

de la distribution des associations par hasard.

Les limites entre les associations de hasard et celles réellement fondées peuvent être estimées à partir des histogrammes précédents. Nous avons choisi de considérer une association réelle si la différence de DM est inférieure à 50 pc.cm^{-3} en même temps qu'une différence de temps inférieure à 0.5 secondes.

La Table 4.4 récapitule les associations avec les événements PRESTO à partir des événements SIPSFAR. Sur un total de 58052 associations, il y en a 18% qui satisfont le critère de proximité choisi ($d\text{DM} < 50 \text{ pc.cm}^{-3}$ et $dt < 0.25 \text{ s}$). Ce faible pourcentage reflète la grande proportion d'associations par hasard révélatrice de seuils de détection bas dans les deux méthodes produisant beaucoup de bruit dans les événements enregistrés. Lorsqu'une sélection des événements SIPSFAR de rapport signal à bruit (SNR) supérieur à 10 est faite, la proportion

Nombre total d'associations	58052
$dDM < 50 \text{ pc.cm}^{-3}$ et $dt < 0.25 \text{ s}$	$10427/58032 = 18 \%$
$SNR > 10$ et $dDM < 50 \text{ pc.cm}^{-3}$, $dt < 0.25 \text{ s}$	$9/28 = 32 \%$

TABLE 4.4 – Associations avec des événements PRESTO à partir des événements SIPSFAR.

d'associations pouvant être jugées réelles augmente comme attendu et s'élève à $\sim 1/3$.

Comme craint à la suite de l'étude détaillée de l'observation 29298000, la comparaison des deux méthodes de détection des événements impulsionnels n'est pas facile et il est difficile de tirer des conclusions plus précises que de dire que la méthode PRESTO est plus sensible que la méthode SIPSFAR, tout en étant bien plus lourde à mettre en œuvre et sensible aux parasites. Il s'agit de savoir quel type de compromis nous sommes prêts à accepter : soit il faut du quasi temps réel mais au prix de la sensibilité, soit on veut de la sensibilité et il faut y passer du temps. En effet, le temps du traitement Presto est de l'ordre de l'heure alors que celui de SIPSFAR est de l'ordre de la seconde pour le traitement d'un fichier de 2 mins.

4.5 Conclusion

Un programme de recherche systématique de nouveaux pulsars a été conduit au radiotélescope de Nançay à la fin des années 1990s avec un échantillonnage temporel rendant possible la détection d'événements impulsifs.

Une recherche d'événements avec la méthode SIPSFAR, basée sur la transformée de Radon a été portée sur GPU et appliquée à l'ensemble des données de ce programme de recherche.

L'analyse d'observations particulières (le pulsar J0248+6021 découvert lors de la recherche initiale) et un traitement statistique simple confirme que la méthode SIPSFAR est moins sensible aux événements faibles tout en étant bien plus rapide en temps de calcul, et bien plus robuste aux parasites.

Dans cette première application, les profils Radon intégrés n'ont pas été exploités. Il est prévu d'étudier son efficacité et de définir pour quel type de pulsar ou de configuration il sera efficace. Il est intéressant de l'exploiter dans l'observation d'amas globulaire qui est un amas stellaire très dense, contenant typiquement une centaine de milliers d'étoiles, dont des pulsars. En effet, vu la grande distance entre la terre et les amas globulaires, et vu que les ondes provenant de ses pulsars traversent pratiquement le même milieu stellaire, ils vont présenter la même mesure de dispersion DM^1 . Plusieurs impulsions de différents pulsars mais de même DM peuvent alors se superposer sur le profil intégré de Radon faisant augmenter le rapport signal sur bruit. Un bon exemple de test éventuel est l'amas globulaire de Terzan 5 de DM compris entre 230 et 250 $pc.cm^{-3}$ et dont on sait déjà qu'il contient plusieurs dizaines de pulsars.

1. <http://www2.naic.edu/pfreire/GCpsr.html>

Conclusion et perspectives

Durant cette thèse nous avons, d'une part étudié des méthodes d'élimination de brouilleurs pour l'observation des pulsars et d'autre part conçu une nouvelle méthode de détection d'impulsions individuelles provenant des pulsars, tout en étant robuste contre les interférences radioélectriques. Les travaux ont été menés en se positionnant dans une optique de traitement en temps réel des données acquises par un radiotélescope à antenne unique. Mais ces méthodes restent valables dans un cadre à antennes multiples comme les réseaux LOFAR et SKA.

La première partie de cette étude a été consacrée à l'investigation des méthodes de traitement des interférences. Leur application dans les observations pulsars nécessite de choisir les plus robustes contre les RFI et les moins sensibles au caractère impulsionnel du signal pulsar. Cette étude montre que le critère cyclostationnaire a un potentiel intéressant dans ce cadre. Mais pour plus d'efficacité, l'utilisation d'autres algorithmes s'appuyant sur des critères différents doit être creusé, voir par exemple ce qui se fait en radio cognitive [Palicot, 2010].

A partir de cet ensemble de détecteurs, nous avons esquissé une architecture potentielle fondée sur la multiplication des points d'analyse. Ceci afin d'avoir accès à différents niveaux de résolution en temps et en fréquence. En effet, les performances des détecteurs de RFI dépendent évidemment de l'adéquation entre l'interférence et les propriétés statistiques (puissance, cyclostationnarité, non gaussianité dans notre cas) qu'ils évaluent. Mais ils dépendent aussi des caractéristiques temporelles et fréquentielles des RFI relativement aux caractéristiques temporelles et fréquentielles du pavage en temps et en fréquence induit par les différents bancs de filtres de la chaîne de réception. Ainsi, une multitude de points d'analyse garantit l'existence d'une résolution en temps et en fréquence amenant une détection optimale du brouilleur. Une fois le status, pollué ou non, d'un bloc de données établi, on peut donner le choix au radioastronome d'exploiter, ou non, ces données. Cette approche est sensée garantir une gestion automatique de RFIs sans risque de perdre des données vitales. Une des options envisagées à Nançay est la mise en oeuvre d'une Base de Données RFI dynamique. Ainsi, les détecteurs répartis tout au long de la chaîne de réception viendront inscrire en temps réel les caractéristiques des événements détectés sur les données en cours d'acquisition. Le radioastronome viendra ensuite consulter cette base de données pour vérifier l'intégrité et la qualité des données qu'il veut exploiter. L'étude que nous avons menée est une étape dans cette direction. Nous y avons ajouté quelques spécificités liées à l'observation des pulsars tels que l'effet d'une impulsion pulsar sur les performances de détection, l'effet de la dédispersion sur ces mêmes performances.

Dans la deuxième partie, nous avons proposé une méthode originale, nommée SIPSFAR, pour la détection des impulsions individuelles de grande puissance en présence de RFI. C'est une méthode de détection aveugle qui n'a besoin de connaître ni le DM et ni la période du pulsar au préalable. Nous avons montré que cette méthode était moins sensible que l'approche classique basée sur une dédispersion incohérente. Cependant SIPSFAR pourrait permettre de dégrossir l'énorme quantité de données acquises lors de la recherche de nouveaux pulsars ou alors être installé en parallèle des observations classiques afin de détecter des événements improbables. Sa robustesse intrinsèque aux RFIs en fait un instrument plutôt sélectif. Une autre utilisation pourrait être l'étude d'amas globulaire. Pour SIPSFAR, une piste d'évolution serait la détection d'impulsions très dispersées. En effet, quand l'impulsion dispersée ne peut plus être considérée comme une droite dans le plan temps fréquence, elle se transforme en un cône dans l'image $2D - FFT$. Ceci bien sûr diminue la sensibilité de SIPSFAR. Dans ce cas, une intégration suivant un cône d'une certaine ouverture pourrait être considérée à la place de l'intégration suivant une droite tel que cela est fait dans la transformée de Radon. Une autre alternative d'amélioration dans le même contexte, est de découper la bande d'observation de sorte à avoir une dispersion linéaire dans les sous bandes obtenues.

D'un point de vue général, il est important de rappeler que l'efficacité de tout traitement des interférences radio est limitée et que la sensibilité d'un radiotélescope est toujours dégradée lorsque de telles techniques doivent être mises en oeuvre. Il ne faut donc pas interpréter ces possibilités de traitement comme des ouvertures

vers un assouplissement des normes d'émission dans les bandes radio. A ce titre, une des pistes à privilégier en priorité est bien évidemment la mise en place de zones de silence radio (Quiet Zone) autour des radiotélescopes.

Annexe A

La radio cognitive

La définition de la radio cognitive adoptée par la Commission Fédérale des Communications (FCC) est un système radio qui prend conscience de son environnement électromagnétique [Palicot, 2010]. Selon le degré de connaissance, il peut ajuster ses paramètres, d'une manière dynamique et autonome, pour modifier son fonctionnement : Maximiser le débit, éviter les interférences, faciliter l'interopérabilité. Ainsi, l'aspect principal de la radio cognitive est lié à l'exploitation autonome du spectre inutilisé par l'utilisateur primaire.

La terminologie de l'utilisateur primaire dans les télécommunications peut être définie comme étant l'utilisateur bénéficiant des droits d'accès à une partie spécifique du spectre. Un utilisateur secondaire par définition ne peut utiliser cette partie du spectre que s'il n'engendre pas d'interférence dans la bande de fréquence réservée à l'utilisateur primaire. Par conséquent, l'utilisateur secondaire a besoin de la radio cognitive pour repérer si le spectre est déjà utilisé ou non, et changer automatiquement ses paramètres radio pour l'exploiter.

L'étape la plus importante et cruciale dans la radio cognitive, est ce qu'on appelle la surveillance du spectre (en anglais spectrum sensing). Elle consiste à mettre en œuvre des algorithmes pour la détection de présence de signaux de télécommunications. Plusieurs approches ont été proposées dans la littérature. Quelques unes sont basées sur la connaissance a priori du signal émis par les utilisateurs primaires, d'autres peuvent détecter en aveugle. Généralement, on s'appuie sur d'autres techniques, en plus des algorithmes de détection, pour minimiser les fausses détections et ainsi mieux exploiter le partage du spectre. Parmi ces techniques on trouve :

- La sauvegarde des informations sur les signaux détectés, comme la fréquence de modulation, le baud rate...
- Le partage des informations entre les systèmes radio cognitive.
- La centralisation des informations qui permettent d'identifier les parties inutilisées du spectre et de les diffuser ensuite aux systèmes radio.
- L'utilisation de l'historique des émissions pour prédire l'avenir.

Le paragraphe suivant recense les méthodes de détection utilisées en radio cognitive. Ensuite, nous nous focaliserons sur les méthodes qui peuvent être envisagées dans le cadre des observations radioastronomiques.

1. Méthodes de surveillance du spectre :

Comme il était expliqué auparavant, la surveillance du spectre consiste à détecter la présence de signaux de télécommunications dans un spectre donné. La problématique est donc exactement la même que celle présentée dans le paragraphe 2.2. Les systèmes de type radio cognitive doivent alors choisir des méthodes se basant sur des critères qui séparent au mieux les deux hypothèses H_0 et H_1 . Pour cette raison, on trouve des méthodes conventionnelles communes à la radio cognitive et à la détection des signaux, telles que :

- Le détecteur de puissance dans la radio cognitive présente les mêmes problèmes que dans la radioastronomie. Les fluctuations du niveau du bruit, la présence des interférences et le canal avec évanouissement sont les principales causes de la chute des performances du détecteur de puissance, de plus il ne peut pas faire la différence entre un utilisateur primaire et un autre secondaire [Cabric et al., 2004]. Sa robustesse peut être améliorée en l'associant à des "empreintes" de l'utilisateur primaire comme le baud rate, le type de modulation ou autre. Dans [Ganesan and Li, 2005] le détecteur de puissance est utilisé en connaissant la puissance du signal émis par l'utilisateur primaire. C'est une sorte de coopération avec les utilisateurs de la bande.
- Des méthodes basées sur la statistique de la covariance ou de l'auto-corrélation du signal reçu ont été proposées dans [Zeng and Liang, 2007]. Les statistiques des matrices de covariance ou d'auto-corrélations du signal et du bruit sont généralement différents. La matrice de covariance statistique

du bruit est déterminée par le filtre de réception. Par conséquent, sa structure est connue par le récepteur, et donc, elle peut être diagonalisée. En conséquence, la matrice de covariance en sortie du processus de diagonalisation est effectivement diagonale si le signal reçu contient uniquement du bruit. Deux méthodes de détection basées sur ce principe sont proposées : La première est la covariance norme 2 (Covariance Absolute Value - CAV) et l'autre est la covariance norme de Frobenius (CFN)[Zeng and Liang, 2007].

- La transformée de Hough a été utilisée dans [Challapali et al., 2004] pour la détection d'un train d'impulsions typique du RADAR.
- Le filtrage adapté est considéré comme la méthode optimum quand on connaît le signal transmis par l'utilisateur primaire. Son avantage principal est le fait qu'il peut atteindre des taux de fausse détection très bas comparé à d'autres méthodes [Tandra and Sahai, 2005]. Par contre, le filtrage adapté dans la radio cognitive nécessite une connaissance exacte du signal émis pour pouvoir le démoduler à la réception. Il faut connaître la bande passante, la fréquence et le type de modulation, la forme d'onde et la forme de la trame. Ceci rend son application très complexe, surtout quand on privilégie des systèmes de radio cognitive capables de détecter la majorité des signaux émis.
- La détection de forme d'onde, autrement appelée détection cohérente, est basée sur la connaissance à priori de la forme d'onde du signal reçu. Plus précisément, la connaissance exacte des trames envoyées par l'émetteur avant chaque transmission. Elles sont utilisées pour des fins de synchronisation entre autres. On trouve dans ce genre de trame la séquence du préambule insérée au début, la séquence d'apprentissage au milieu, etc, qui sont en général connues. En conséquence, la détection peut se faire en corrélant le signal reçu avec une version antérieure connue [Tang, 2005] [Sahai et al., 2006] [Urkowitz, 1967]. Cette méthode est applicable uniquement lorsqu'on connaît les séquences qui forment le signal utile. Les résultats présentés dans [Tang, 2005] montrent que cette méthode est plus performante que le détecteur de puissance. En plus, les résultats montrent que les performances augmentent avec la longueur de la séquence connue. Des tests ont été faits avec des signaux de séquences de la norme IEEE 802.11b [Sup, 1999] dans [Geirhofer et al., 2006a] et [Geirhofer et al., 2006b] et démontrent que le temps de convergence de la détection cohérente est court ; par contre elle est sensible à des problèmes de synchronisation.
- Le test multi-cyclique est une méthode basée sur le principe de cyclostationnarité. L'algorithme multi-cyclique a été détaillé dans [Ghozzi et al., 2006], il permet la recherche de la cyclostationnarité pour plusieurs fréquences cycliques. Les tests ont été effectués pour la détection des canaux de télévision libres (Les systèmes de SECAM, PAL, NTSC et DVBT TV). Les résultats montrent que cette méthode est plus performante que le détecteur de puissance quand on suppose que la densité spectrale de puissance du bruit n'est pas stable, ce qui est une supposition toute à fait réaliste. Le critère de cyclostationnarité est aussi utilisé dans [Shankar et al., 2005]. La plupart des méthodes citées ci-dessus requièrent des connaissances à priori du signal émis par l'utilisateur primaire afin de le détecter

Annexe B

Impact du filtre de mise en forme

Le filtre de mise en forme, $h(t)$ est celui qui, dans une chaine d'émission de signaux de télécommunication, donne le gabarit de la réponse spectrale du signal. Il genere également la forme temporel des bits émis. Les filtres classiquement utilisés sont les filtres rectangulaires, les filtres gaussiens et les filtres en cosinus surelevé ou racine de cosinus surélevé.

Dans le cadre des calculs des corrélations cycliques et cycliques conjuguées, ce filtre intervient dans le calcul des amplitudes de ces corrélations (Equ. 2.9 et Equ. 2.10) sous la forme :

$$r_{h,h^{(*)}}^{\frac{l}{T_{sym}}}(\tau) = \int_{-\infty}^{+\infty} h(t + \frac{\tau}{2})h^{(*)}(t - \frac{\tau}{2})e^{-j2\pi\frac{l}{T_{sym}}t} dt \quad (B.1)$$

où $\frac{l}{T_{sym}}$ est le baud rate.

Pour augmenter l'information cyclostationnaire disponible, il est important donc de maximiser l'amplitude de cette fonction. Dans nos applications, nous avons considéré le cas $\tau = 0$, mais parfois l'optimum se trouve pour des valeurs $\tau \neq 0$. C'est ce que nous allons montrer dans la suite.

Nous allons donc comparer le cas où $h(t)$ est un filtre de type « racine de cosinus surélevé » (rrcos) et le cas où il est « rectangulaire ». La figure B.1 présente la réponse impulsionnelle de ces deux filtres ainsi que la fonction $r_{h,h^{(*)}}^{\alpha}$ associée, en fonction de α . Les cercles rouges correspondent aux valeurs de $\alpha = l/T_{sym}$ avec $T_{sym} = 8$ et $l = 0, \dots, 8$.

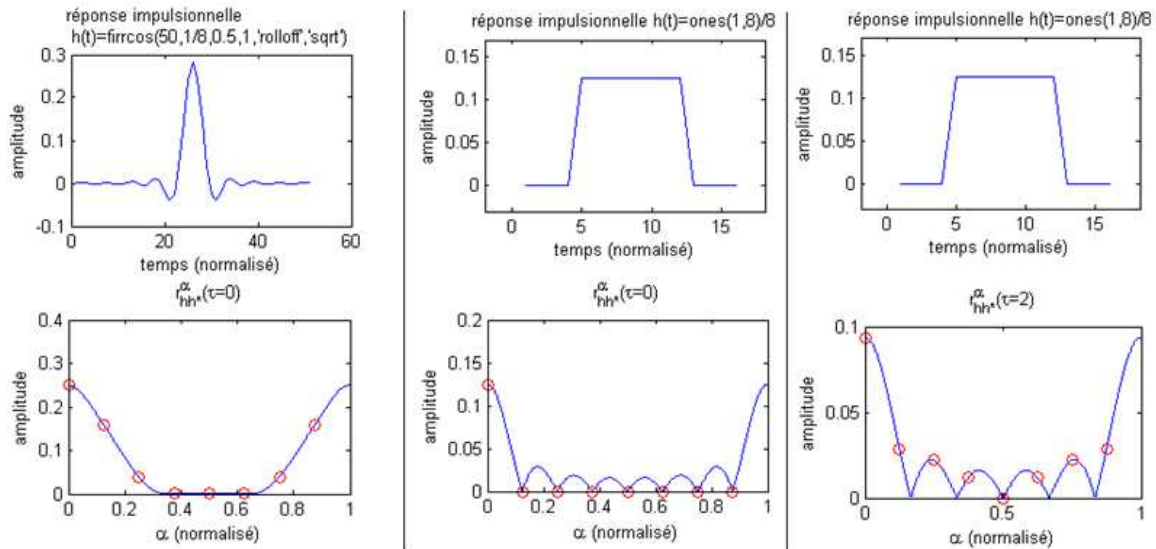


FIGURE B.1 – Réponses impulsionnelles $h(t)$ (haut) et $r_{h,h^{(*)}}^{\alpha}$ (bas). Les cercles rouges correspondent aux positions des fréquences cycliques liées au baud rate de la modulation, ici $l/8$ avec $l = 0, \dots, 8$. On remarque que dans le cas d'un filtre rectangulaire, avec $\tau = 0$ seule la fréquence $\alpha = 0$ donne une amplitude non nulle.

Pour la corrélation cyclique, $r_{h,h^{(*)}}^{\alpha}$ peut être calculée sur les valeurs $\alpha = l/T_{sym}$ avec $l = 1, \dots, 8$ alors que pour la corrélation cyclique conjuguée, le cas $l = 0$ est également possible. Or c'est pour ce dernier cas que le

maximum de r_{h,h^*}^α est obtenu. C'est ce qui explique l'écart de performances entre le détecteur cyclique et le détecteur cyclique conjugué dans les simulations du chapitre 2.

Avec le cas de la fenêtre rectangulaire, et pour $\tau = 0$ (courbe du milieu), il n'y a pas d'information cyclique (sauf pour la corrélation cyclique conjuguée à cause du cas $l = 0$). Donc le seul moyen d'obtenir de l'information dans le cas cyclique est de prendre un τ différent de 0 (courbe de droite où on retrouve des valeurs non nulles pour $l = 1, \dots, 8$).

Annexe C

La transformée de Radon

– **Définition de la transformée de Radon :**

La transformée de Radon a été définie de plusieurs façons dans la bibliographie et selon les applications. En sismique [Yilmaz and Doherty, 1987][Deans, 1983], elle était définie comme étant des intégrations inclinées (Slant Stacking) donnant l'emplacement d'une ligne dans l'image initiale par sa pente p et son offset τ . La transformée de Radon $\check{g}(p, \tau)$ d'une fonction continue $g(x, y)$ est alors l'intégration de g le long de lignes inclinées et est donnée par :

$$\check{g}(p, \tau) = \int_{-\infty}^{+\infty} g(x, px + \tau) dx \quad (\text{C.1})$$

Une autre définition adoptée dans d'autres disciplines comme la tomographie, l'astronomie et la microscopie [Deans, 1983], où on décrit une ligne en coordonnées polaires ρ et θ comme suit :

$$\rho = x \cos \theta + y \sin \theta \quad (\text{C.2})$$

La transformée de Radon est donc donnée par :

$$\check{g}(\rho, \theta) = \int_{-\infty}^{+\infty} g(\rho \cos \theta - s \sin \theta, \rho \sin \theta + s \cos \theta) ds \quad (\text{C.3})$$

Avec s est l'axe d'intégration.

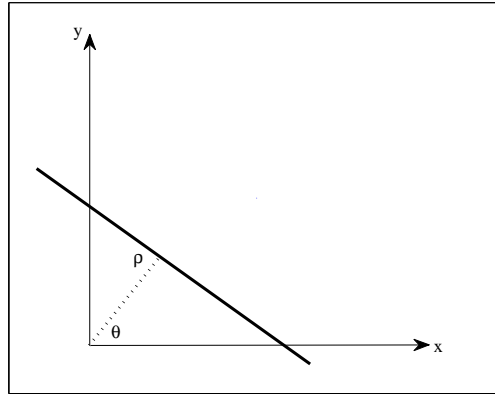


FIGURE C.1 – La position d'une ligne en coordonnées polaires

– **La transformée de Radon d'une droite :**

Dans ce manuscrit, on s'intéresse uniquement à la détection d'une droite (ligne) dans une image noyée dans un bruit gaussien. Pour ce fait, on va démontrer ici que la transformée de Radon d'une droite est un

pic. Considérons une droite en deux dimensions ayant comme paramètres (ρ_1, θ_1) pouvant être modélisé par la fonction de Dirac suivante :

$$g(x, y) = \delta(\rho_1 - x \cos \theta_1 - y \sin \theta_1) \quad (\text{C.4})$$

Donc :

$$\begin{aligned} \check{g}(\rho, \theta) &= \int_{-\infty}^{+\infty} \delta(\rho_1 - (\rho \cos \theta - s \sin \theta) \cos \theta_1 - (\rho \sin \theta - s \cos \theta) \sin \theta_1) ds \\ &= \int_{-\infty}^{+\infty} \delta(\rho_1 - \rho \cos(\theta - \theta_1) + s \sin(\theta - \theta_1)) ds \\ &= \int_{-\infty}^{+\infty} \frac{1}{|\sin(\theta - \theta_1)|} \delta\left(\frac{\rho_1 - \rho \cos(\theta - \theta_1)}{\sin(\theta - \theta_1)} + s\right) ds \quad \text{si } \sin(\theta - \theta_1) \neq 0 \\ &= \frac{1}{|\sin(\theta - \theta_1)|} \end{aligned} \quad (\text{C.5})$$

Et si $\theta = \theta_1$, c-à-d. $\sin(\theta - \theta_1) = 0$, on trouve que :

$$\check{g}(\rho, \theta) = \int_{-\infty}^{+\infty} \delta(\rho_1 - \rho) ds = \begin{cases} 0 & \text{si } \rho_1 \neq \rho \\ \int_{-\infty}^{+\infty} \delta(0) ds & \text{si } \rho_1 = \rho \end{cases} \quad (\text{C.6})$$

On en déduit que le résultat est un pic formé quand $\rho = \rho_1$ et $\theta = \theta_1$. Plus de détails peuvent être trouvés dans [Toft, 1996a].

– **La transformée de Radon discrète** : On considère les variables discrètes suivant :

$$\begin{aligned} x &= x_m = x_{min} + m\Delta x, \quad m = 0, 1, \dots, M-1 \\ y &= y_n = y_{min} + n\Delta y, \quad n = 0, 1, \dots, N-1 \\ \theta &= \theta_l = \theta_{min} + l\Delta\theta, \quad l = 0, 1, \dots, L-1 \\ \rho &= \rho_k = \rho_{min} + k\Delta\rho, \quad k = 0, 1, \dots, K-1 \end{aligned} \quad (\text{C.7})$$

Où (x_{min}, y_{min}) est la position du premier pixel, Δx est le pas d'échantillonnage de x , et m est l'indice de l'échantillon pour la variable x . Les mêmes notations sont utilisées pour la variable y . Pour minimiser le nombre des échantillons ρ_k [Toft, 1996a], il était convenu de replacer l'origine des pixels ($x = 0, y = 0$) au centre de l'image $g(x, y)$. On suppose maintenant que l'image est carrée :

$$\Delta x = \Delta y \quad (\text{C.8})$$

$$M = N \quad (\text{C.9})$$

Ce qui implique des échantillons symétriques par rapport à l'origine :

$$x_{min} = -x_{max} = -\frac{M-1}{2}\Delta x \quad (\text{C.10})$$

$$y_{min} = -y_{max} = -\frac{M-1}{2}\Delta y \quad (\text{C.11})$$

$$\rho_{min} = -\rho_{max} = -\frac{K-1}{2}\Delta\rho \quad (\text{C.12})$$

Radon doit balayer l'ensemble des θ entre $[0, \pi]$ avec un pas que l'utilisateur peut choisir.

Une façon pour définir la transformée de Radon discrète (Discret Radon Transform DRT) pouvait être en remplaçant les variables continues par les variables discrètes données par l'équation C.7 dans l'équation C.3, et l'intégrale continue par une somme. La transformée de Radon discrète peut être alors donnée par :

$$\check{g}(\rho_k, \theta_l) \approx \Delta s \sum_{j=0}^{J-1} g(\rho_k \cos \theta_l - s_j \sin \theta_l, \rho_k \sin \theta_l + s_j \cos \theta_l) \quad (\text{C.13})$$

Où s_j est un échantillon de la variable s . Cette approximation présente a un point faible. En effet, pour des valeurs de j , le couple $(x_m, y_n) = (\rho_k \cos \theta_l - s_j \sin \theta_l, \rho_k \sin \theta_l + s_j \cos \theta_k)$ ne coïncide pas avec des pixels de l'image $g(m, n) = g(x_m, y_n)$, d'où la nécessité de faire une interpolation en deux dimensions. Plusieurs méthodes peuvent être utilisées dans ce cas là, interpolation linéaire, interpolation avec le plus proche voisin ou l'interpolation en sinus cardinal. L'équation C.14 présente ici la transformée de Radon discrète avec une interpolation du plus proche voisin qui est la plus adaptée dans la détection d'une droite dans l'image [Toft, 1996b][Hansen and Toft, 1996][Toft, 1996a].

$$\check{g}(\rho, \theta) \approx \begin{cases} \frac{\Delta x}{|\sin \theta|} \sum_{m=0}^{M-1} g(m, \left\lceil -m \cot \theta + \frac{\rho - x_{\min}(\cos \theta + \sin \theta)}{\Delta x \sin \theta} \right\rceil) & \text{si } \sin \theta > \frac{1}{\sqrt{2}} \\ \frac{\Delta x}{|\cos \theta|} \sum_{n=0}^{M-1} g(\left\lceil -n \tan \theta + \frac{\rho - x_{\min}(\cos \theta + \sin \theta)}{\Delta x \cos \theta} \right\rceil, n) & \text{si } \sin \theta \leq \frac{1}{\sqrt{2}} \end{cases} \quad (\text{C.14})$$

Où $M = N$, $\Delta x = \Delta y$, et $x_{\min} = y_{\min}$.

Annexe D

Calcul des sensibilités $S_{SIPSFAR}$ et S_{PRESTO}

– Les statistiques du plan T-F :

Le plan T-F en puissance est calculée en effectuant des FFT sur des blocs N successifs. Ceci peut être vu comme un banc de filtre avec une fenêtre de pondération rectangulaire qui répartie l'énergie totale du signal sur les sous- bandes de fréquence. La réponse impulsionnelle de ce filtre dans ce cas est :

$$\sum_{i=0}^{N-1} h(i) = \frac{1}{N} \quad (D.1)$$

$H(f)$ est la réponse fréquentielle de $h(n)$. Considérons le signal $x(n)$ qui est un bruit blanc complexe gaussien centré de variance σ_x^2 (i.e. $\mathcal{N}(0, \sigma_x^2)$). $X(f)$ est le spectre de $x(n)$. Soit $|X_k(f)|^2$ sa sortie du banc de filtre pour un canal k , et en utilisant l'égalité de Parseval on a alors :

$$\begin{aligned} \int_0^1 |X_k(f)|^2 df &= \int_0^1 |X(f)|^2 |H(f)|^2 df \\ &= \sigma_x^2 \int_0^1 |H(f)|^2 df \\ &= \sigma_x^2 \sum_{n=0}^{N-1} |h(n)|^2 \\ &= \frac{\sigma_x^2}{N} \end{aligned} \quad (D.2)$$

Le filtrage étant une opération linéaire, la variable $X_k(f)$ dans un canal, où il y a uniquement la contribution du bruit, suit une loi normale centrée de variance $\frac{\sigma_b^2}{N}$. De même pour l'endroit où on a le pulse, son maximum est de l'ordre de $\frac{\sigma_p^2}{N} + \frac{\sigma_b^2}{N}$.

Nous allons maintenant faire le calcul de la moyenne et de la variance de $|X_k(f)|^2 = I^2 + R^2$ la puissance instantanée du signal dans un canal k , avec $(I, R) \propto \mathcal{N}(0, \frac{\sigma_b^2}{2N})$.

Considérons la variable aléatoire Y qui suit une loi de χ^2 centrée de 2 degrés de liberté de moyenne égale à 2 et de variance égale à 4 :

$$Y = \left(\frac{a + \mu_a}{\sigma_a} \right)^2 + \left(\frac{b + \mu_b}{\sigma_b} \right)^2 \quad (D.3)$$

La moyenne de $|X_k(f)|^2$ peut être calculée alors :

$$\begin{aligned} E[Y] &= E \left[2N \frac{a^2}{\sigma_a^2} + 2N \frac{b^2}{\sigma_b^2} \right] \\ E[Y] &= \frac{2N}{\sigma_b^2} E[a^2 + b^2] \\ 2 &= \frac{2N}{\sigma_b^2} E[|X|^2] \end{aligned} \quad (D.4)$$

La moyenne de $|X_k(f)|$ est donc :

$$E[|X|^2] = \frac{\sigma_b^2}{N} \quad (D.5)$$

De même pour la variance, on a :

$$\begin{aligned} var(Y) &= \frac{4N^2}{\sigma_b^4} E[|X_k(f)|^4] - \frac{4N^2}{\sigma_b^4} E^2[|X_k(f)|^2] \\ 4 &= \frac{4N^2}{\sigma_b^4} var(|X_f|^2) \end{aligned} \quad (D.6)$$

Donc la variance de la variable $|X_k(f)|^2$ est :

$$var(|X_f|^2) = \frac{\sigma_b^4}{N^2} \quad (D.7)$$

La puissance instantanée $|X_k(f)|^2$, du signal $x(t)$ dans chaque canal k , suit la loi de χ^2 de moyenne $\frac{\sigma_b^2}{N}$ et de variance $\frac{\sigma_b^4}{N^2}$. Dans ce qui suit, nous utilisons les corollaires suivant avec μ la moyenne statistique :

Corollaire 1 : La transformée de Fourier normalisée (on divise par N) d'une variable complexe aléatoire $\propto \mathcal{N}(\mu, \sigma^2)$ est une variable complexe aléatoire $\propto \mathcal{N}(\frac{\mu}{N}, \frac{\sigma^2}{N})$.

Corollaire 2 : La transformée de Fourier d'une variable complexe aléatoire de moyenne μ et de variance σ^2 est un signal complexe aléatoire de moyenne $N\mu$ et de variance $N\sigma^2$.

Corollaire 3 : Le module au carré d'un signal complexe aléatoire $\propto \mathcal{N}(\mu, \sigma^2)$ est un signal aléatoire réel $\propto \chi^2(\mu + \sigma^2, \sigma^4)$.

Nous en déduisons que les échantillons où l'on a la contribution du pulsar et le bruit dans le plan temps-fréquence suivent la loi de $\chi^2(\frac{\sigma_a^2}{N}, \frac{\sigma_b^4}{N^2})$, avec $\sigma_a^2 = \sigma_p^2 + \sigma_b^2$

- **Les statistiques du signal en sortie de SIPSFAR :** La transformée de Fourier à deux dimensions appliquée au spectre dynamique peut être divisée en deux étapes ; une première TF sur les lignes suivit par une deuxième sur les colonnes.

A la sortie de la TF sur les lignes, la puissance du bruit est :

$$P_b^l(k) = \begin{cases} \sigma_b^2 + \frac{\sigma_b^4}{N} & \text{si } k = 0, \\ \frac{\sigma_b^4}{N} & \text{si non} \end{cases} \quad (D.8)$$

A la sortie de la TF sur les colonnes, la puissance du bruit est :

$$P_b^c(k) = \begin{cases} N\sigma_b^2 + \sigma_b^4 & \text{si } k = 0, \\ \sigma_b^4 & \text{si non} \end{cases} \quad (D.9)$$

Le module au carré du bruit en sortie de la $2D - FFT$ est :

$$P_b(k) = \begin{cases} (N^2 + 1)\sigma_b^4 + \sigma_b^8 & \text{si } k = 0, \\ \sigma_b^4 + \sigma_b^8 & \text{si non} \end{cases} \quad (D.10)$$

Après intégration sur N canaux (N lignes de l'image obtenue) et sur M images, le bruit de ce détecteur a les statistiques suivantes :

$$\mu_{SIPSFAR} = \begin{cases} (N + 1)\sigma_b^4 & \text{si } k = 0, \\ \sigma_b^4 & \text{si non} \end{cases} \quad (D.11)$$

$$\sigma_{SIPSFAR} = \frac{\sigma_b^4}{\sqrt{NM}} \quad \forall k$$

Avec $\mu_{SIPSFAR}$ et $\sigma_{SIPSFAR}$ sont respectivement la moyenne et l'écart type de SIPSFAR. La même méthodologie est suivie pour le calcul de la puissance du pulsar en sortie du détecteur, sans oublier que le signal pulsar est multiplié par une porte rectangulaire de longueur L (c'est la forme du pulse considérée ici), et donc après la TF sur les colonnes, un sinus cardinal va apparaître dans le canal zéro :

$$P_p = T\sigma_p^2 \text{sinc}(\pi Tl) \quad (\text{D.12})$$

Et donc, en intégrant sur N canaux, le pic du pulsar plus le bruit a comme amplitude :

$$A = T^2\sigma_p^4 + (N+1)\sigma_b^4 \quad (\text{D.13})$$

D'où d qui la distance entre le pic de l'éventuel pulsar et la moyenne du bruit du détecteur comme il est montré dans la figure 3.6(b) :

$$\begin{aligned} d &= A - \mu_{SIPSFAR} \\ d &= T^2\sigma_p^4 \end{aligned} \quad (\text{D.14})$$

- **Les statistiques du signal en sortie de PRESTO** : Vu qu'on est dans le cas du $DM = 0$, il suffit de faire une intégration sur les N canaux et ensuite sur M profils. Les statistiques du détecteur sont :

$$\begin{aligned} \mu_{PRESTO} &= \sigma_b^2 \\ \sigma_{PRESTO} &= \frac{\sigma_b^2}{\sqrt{NM}} \end{aligned} \quad (\text{D.15})$$

De même l'amplitude de l'impulsion pulsar est :

$$A = \sigma_p^2 + \sigma_b^2 \quad (\text{D.16})$$

et donc :

$$d = \sigma_p^2 \quad (\text{D.17})$$

La sensibilité de PRESTO est alors :

$$S_{PRESTO} = \sqrt{NMSNR} \quad (\text{D.18})$$

Annexe E

Code source du traitement SIPSFAR

```
#include <stdlib.h>
#include <stdio.h>
#include <string.h>
#include <time.h>
#include "fbreadinf.c"
#include "jdgregjdc.c"
#include "dateheure.c"
#include <cuda.h>
#include <cuda_runtime.h>
#include <cutil_inline.h>
#include <cutil.h>
#include <calc_sens.h>
#include <calc_sens.c>
#define MaxT 1024*1024

// size of the matrix
#define threads_per_block 128
// block size
#define blocks_per_grid_row 128

typedef float2 Complex;

void calc_index( int M, int N, int *index );
void shift_index( int M, int N, int *idx );
__global__ void radon( float* img_In, float* radonResults, int* index, int N, int M );
__global__ void sqr_abs( cufftComplex * img_In, float* Result, int M );
__global__ void cum_sum( float * vect, float* result, int M );
__global__ void cu_fftshift( float* img_In, float* img_Out, int* idx, int N, int M );

void swapx4(char *ptr) // flip bits right to left
{
    char tmp;
    tmp = ptr[0];
    ptr[0] = ptr[3];
    ptr[3] = tmp;
    tmp = ptr[1];
    ptr[1] = ptr[2];
    ptr[2] = tmp;
}

int main(int argc, char *argv[])
{
    FILE *pf1;
    int jnbbinav, jrec, size, sizebloc, sizeheader;
    unsigned int rawdata[98304], header[8192], i4d;
    float tr[6], fg, bg, fgf=1024., bgf=1024., fl, xmin, xmax, ymin, ymax;
    long long int i8d;
    float dm;
    double tdep, tend, t, jd0, jd0f, foa, samprate, rh, rm, rs, lengthofintegration, h0, h, h1, h2;
    double Freqs[96], FrOrdered[96], TickOffset, Latency, dt, FrCent, BwChan, ActualTotBW;
    int JFrOrder[96], JPointer[96], id;
    char grdev[80], filename[80], ifilename[80], RACH[32], DECCH[32], StartCH[32], label[80];
    char date[20], date0[16], heure[10], heure0[16], ts[12], *adr;
    short i2d;
    int jnbt, i, j, k, kk, it, iti, jj, j1, j2, k1, k2, jy, jm, jd, jh, jmn, js;
    int idim, jdim, JLAG[96], jrad, jram, jras, jdec, jdec, jdec;
    int ico, chanout[96];
    char st1[80], st2[80];
    float * data;
    cufftComplex * data_padded;
    data_padded = (cufftComplex*) malloc(128*MaxT*sizeof(cufftComplex));
```

```

data = (float*)malloc(96*MaxT*sizeof(float));
printf("size of data %d\n",96*MaxT*sizeof(float));
// CUFFT 2D plan
cufftHandle plan;
cufftPlan2d(&plan, 128, 128, CUFFT_C2C);
FILE *fp;
char NOM_FICHIER[256];
float *d_radon,*sum_radon,*d_abs_fft_data,*d_abs_fft_data_clean;
float *c_radon,*c_abs_fft_data,*radon_cum;
int *d_index,*d_idx;
cufftComplex *d_data,*c_fft_data, *d_fft_data;
c_fft_data = (cufftComplex*)malloc(128*128*sizeof(cufftComplex));
c_abs_fft_data=(float*)malloc(128*128*sizeof(float));
c_radon=(float*)malloc(128*sizeof(float));
radon_cum=(float*)malloc(128*sizeof(float));
int dataAmount = threads_per_block * blocks_per_grid_row;
int blockGridWidth = blocks_per_grid_row; //
int blockGridHeight = (dataAmount / threads_per_block) / blockGridWidth; //
dim3 blockGridRows( blockGridWidth, blockGridHeight );//
dim3 threadBlockRows(threads_per_block, 1);//
// calculate index to be used by Radon
int * h_index, *h_idx;
h_index = (int*)malloc(128*128*sizeof(int));
h_idx = (int*)malloc(128*128*sizeof(int));
for(int i=0;i<128*128;i++) h_index[i]=0;
for(int i=0;i<128*128;i++) h_idx[i]=0;
calc_index(128, 128, h_index);
shift_index( 128, 128, h_idx);
if(argc>=2)
{
    strcpy(filename,argv[1]);
    sprintf(grdev,"%x");
    dm=0.0;
    jnbbinav=128; // number of time samples binned together
    for(i=2;i<argc;i++)
    {
        if(strncmp(argv[i],"grdev=",6)==0) sscanf(argv[i],"grdev=%s",grdev);
        if(strncmp(argv[i],"bin=",4)==0) sscanf(argv[i],"bin=%d",&jnbbinav);
    }
}
else
{
    fprintf(stderr," Use : %s <FileName> [bin=<JnbBinAv>]\n",argv[0]);
    printf("Will look for Giant pulses using SIPSFAR after crude integration (binning).\n");
    printf("FileName: .fb file name (a .i file must exist)\n");
    printf("JnbBinAv: number of time samples binned together, default is 128\n");
    exit(2);
}

dateheureimpr(date,heure,ts);
sprintf(label,"fbplotc : %s %s - file=%s",date,heure,filename);

strcpy(ifilename,filename);
printf("%s> Reading .i file\n",argv[0]);
fbReadInf(ifilename, Freqs, FrOrdered, JFrOrder, JPointer, ...
RACH, DECCH, &id, StartCH, &TickOffset, &dt, &FrCent, &BwChan);
printf("%s> RA=%s DEC=%s\n",argv[0],RACH,DECCH);
printf("%s> id=%d dt=%lf FrCent=%lf BwChan=%lf\n",argv[0],id,dt,FrCent,BwChan);

/* Extract the PULSAR name */
sscanf(RACH,"%d:%d:%d",&jrad,&jram,&jras);
sscanf(DECCH,"%d:%d:%d",&jdec,&jdec,&jdecs);

sizebloc=sizeof(rawdata); sizeheader=sizeof(header);
printf("%s> sizebloc=%d\n",argv[0],sizebloc); printf("%s> sizeheader=%d\n",argv[0],sizeheader);
printf("%s> sizeof(long)=%d\n",argv[0],sizeof(long));
printf("%s> sizeof(float)=%d\n",argv[0],sizeof(float));
printf("%s> sizeof(short)=%d\n",argv[0],sizeof(short));

/*Start reading file*/
iti=jnbbinav; // jnbbinav = bin
t=0.0e0;
if((pfl=fopen(filename,"r")) != NULL)
{
    fread(header,sizeheader,1,pfl); int count = 0;
    while(!feof(pfl))
    {
        for(i=0;i<96;i++) for(j=0;j<MaxT;j++) data[i+j*96]=0.0e0;
        jrec=0; size=0; tdep=t; it=0;
        while((fread(rawdata,sizebloc,1,pfl) != 0) && (size < 48*MaxT))
        {
            jrec++; size+=sizebloc;
            printf("read bloc # %d size=%d t=%lf it=%d\n",jrec,size,t,it);*/
            k=0;
            for(i=0;i<8192;i++)

```

```

    {
        kk=k;
        jj=0;
        for(j=0;j<12;j++)
        {
            i4d=rawdata[kk];
            #if IS_LITTLE_ENDIAN
                swapx4((char *)&i4d);
            #endif
            i8d=i4d;
            if(i8d < 0) { printf("i8d=%lld ", i8d); i8d=i8d+4294967296; printf("i8d=%lld\n", i8d);}
            i2d=(i8d>>28) & 15; data[JPointer[jj++]+it*96]+=(float)i2d; //+16 for padding
            i2d=(i8d>>24) & 15; data[JPointer[jj++]+it*96]+=(float)i2d;
            i2d=(i8d>>20) & 15; data[JPointer[jj++]+it*96]+=(float)i2d;
            i2d=(i8d>>16) & 15; data[JPointer[jj++]+it*96]+=(float)i2d;
            i2d=(i8d>>12) & 15; data[JPointer[jj++]+it*96]+=(float)i2d;
            i2d=(i8d>>8) & 15; data[JPointer[jj++]+it*96]+=(float)i2d;
            i2d=(i8d>>4) & 15; data[JPointer[jj++]+it*96]+=(float)i2d;
            i2d=i8d & 15; data[JPointer[jj++]+it*96]+=(float)i2d;
            kk++;
        } /* LOOP for(j=0;j<12;j++) */
        iti--;
        if(iti <=0) { it++; iti=jnbbinav; }
        k+=12;
        t+=dt;
    } /* LOOP for(i=0;i<8192;i++) */
} /* LOOP while((fread(rawdata, sizebloc, 1, pf1) != 0) && (size < 48*1024*1024)) */

jnbt = it;
printf("jnbt=%d t=%lf\n", jnbt, t);

/* Normalisation between 0.0 and 1.0 */
for(i=0;i<96;i++)
{
    fg=data[i];
    bg=data[i];
    for(j=0;j<jnbt;j++)
    {
        if(fg > data[i+j*96]) fg=data[i+j*96];
        if(bg < data[i+j*96]) bg=data[i+j*96];
    }

    if(fg != bg)
    {
        for(j=0;j<jnbt;j++) data[i+j*96]=(data[i+j*96]-fg)/(bg-fg);
    }
    else
    {
        for(j=0;j<jnbt;j++) data[i+j*96]=0.0;
    }
}

/* padding chennels */
for(i=0;i<128;i++) for(j=0;j<jnbt;j++)
{
    data_padded[i+j*128].x = 0.0e0;
    data_padded[i+j*128].y = 0.0e0;
}
for(i=16;i<96+16;i++)
{
    for(j=0;j<jnbt;j++)
    {
        data_padded[i+j*128].x=data[i-16+j*96];
    }
}

/***** Data processing in the CPU for comparing *****/
hl=dateheure(date, heure);
fftw_complex *in, *out;
fftw_plan p;
in = (fftw_complex *) fftw_malloc(sizeof(fftw_complex)*128*jnbt);
out = (fftw_complex *) fftw_malloc(sizeof(fftw_complex)*128*128);
abs_out = (float *) malloc(sizeof(float)*128*128);
for(i=0;i<128;i++) for(j=0;j<jnbt;j++){
    in[i+j*128].x=data_padded[i+j*128].x;
    in[i+j*128].y=data_padded[i+j*128].y;}

for(int nbimg=0; nbimg<(jnbt/128); nbimg++)
{
    p = fftw_plan_dft_2d(128, 128, (fftw_complex *) (in+128*128*nbimg), ...
        out, FFTW_FORWARD, FFTW_ESTIMATE);
    fftw_execute(p);
    for(int i=0; i<128; i++) for(int j=0; j<128; j++){
        abs_out[i+j*128] = (out[i+j*128].x*out[i+j*128].x)+(out[i+j*128].y*out[i+j*128].y);}
}

```

```

float * U;
U = (float *) malloc (128 * sizeof(float));
int idx;
for (int i=0; i<128; i++)
{U[i]=0;
for (int j=0; j<128; j++)
{
idx = h_index[128*i+j];
U[i]+= abs_out[idx];
}}
}
h2=dateheure(date0,heure0)-h1; printf("Processing time in CPU= %.6lf sec\n",h2);*/

/*****
/***** Data processing in the GPU *****/
int sizec = 128*jnbt*sizeof(cufftComplex);
float *test;
test = (float*)malloc(sizeof(float)*128);
h0=dateheure(date0,heure0);
cutilSafeCall(cudaMalloc( (void*)&d_data, sizec ));
cutilSafeCall(cudaMalloc( (void*)&d_index, 128*128*sizeof(int) ));
cutilSafeCall(cudaMalloc( (void*)&d_idx, 128*128*sizeof(int) ));
cutilSafeCall(cudaMalloc( (void*)&d_radon, 128*128*sizeof(float) ));
cutilSafeCall(cudaMalloc( (void*)&sum_radon, 128*sizeof(float) ));
cutilSafeCall(cudaMalloc( (void*)&d_fft_data, 128*128*sizeof(cufftComplex) ));
cutilSafeCall(cudaMalloc( (void*)&d_abs_fft_data, 128*128*sizeof(float) ));
cutilSafeCall(cudaMalloc( (void*)&d_abs_fft_data_clean, 128*128*sizeof(float) ));
cutilSafeCall(cudaMemcpy( d_data, data_padded, sizec, cudaMemcpyHostToDevice ));
cutilSafeCall(cudaMemcpy( d_index, h_index, 128*128*sizeof(int), cudaMemcpyHostToDevice ));
for (i=0;i<128;i++){radon_cum[i]=0.0;}
cutilSafeCall(cudaMemcpy( sum_radon, radon_cum, 128*sizeof(float) , cudaMemcpyHostToDevice ));
int nbimg;
//Begin loop on number of images nbimg (8192 = 1024*1024/128)
for (nbimg=0; nbimg<(jnbt/128); nbimg++)
{
// FFT2D 128x128
cufftSafeCall( cufftExecC2C( plan, (cufftComplex *) (d_data+128*128*nbimg), ...
(cufftComplex *) d_fft_data, CUFFT_FORWARD));
cutilSafeCall(cudaMemcpy( c_fft_data, d_fft_data, 128*128*sizeof(cufftComplex), ...
cudaMemcpyDeviceToHost ));
// |d_fft_data|^2
sqr_abs<<<blockGridRows, threadBlockRows>>>(d_fft_data, d_abs_fft_data, 128);
// Check if kernel execution generated an error
cutilCheckMsg("Kernel execution failed [ ComplexPointwiseMulAndScale ]");
cutilSafeCall( cudaThreadSynchronize() );
// RFI blanking
cutilSafeCall(cudaMemcpy( d_idx, h_idx, 128*128*sizeof(int), cudaMemcpyHostToDevice ));
cu_fftshift<<<blockGridRows, threadBlockRows>>>(d_abs_fft_data, d_abs_fft_data_clean, ...
d_idx, 128, 128);
// Radon each FFT2D and save them
radon<<<blockGridRows, threadBlockRows>>>(d_abs_fft_data_clean, d_radon, d_index, 128, 128);
// Check if kernel execution generated an error
cutilCheckMsg("Kernel execution failed [ ComplexPointwiseMulAndScale ]");
cutilSafeCall( cudaThreadSynchronize() );
// Sum all the Radon and save it
cum_sum<<<blockGridRows, threadBlockRows>>>(d_radon, sum_radon, 128);
// if detect Pulse, save d_radon
cutilSafeCall(cudaMemcpy( c_radon, d_radon, 128*sizeof(float) , cudaMemcpyDeviceToHost ));
float theta[128];
float pas=180.00/128, t1, t2;
float max=(float)(pas*128);
theta[0]=0;
for (i=1; i<128;i++){
theta[i]=theta[i-1]+pas;}

float S=0.0, pos=0.0; // S: Sensitivity; pos: theta at max
calc_sens(theta, c_radon, 128, &S, &pos);
dm=tanf((90.0-pos)*3.14159265/180.0)*810.586*jnbbinav/64;
if (S>5.5)
{
printf("----> DETECTION AT TETHA = %f° DM = %f WITH SNR = %f, IN IMAGE NUMBER %d \n",
pos, dm, S, nbimg+count);
t1 = (nbimg+count)*127*jnbbinav*dt;
t2 = (nbimg+count+1)*127*jnbbinav*dt;
sprintf(NOM_FICHIER, "GPDet-radon_%08d_%03d-%06d-%04d-%03d.dat",
id, jnbbinav, (int)(1000.0*t1), (int)dm, (int)S);
fp = fopen(NOM_FICHIER, "w");
fprintf(fp, "# Scan = %d Tetha = %f° DM= %f SNR = %f in time window = [%f %f] s \n",
id, pos, dm, S, t1, t2);
for (i = 0; i < 128 ; i++){ fprintf(fp, "%f \n", c_radon[i]);}
fclose(fp);
}
// Check if kernel execution generated an error
cutilCheckMsg("Kernel execution failed [ ComplexPointwiseMulAndScale ]");
cutilSafeCall( cudaThreadSynchronize() );

```

```

    } // End loop
    count=nbimg;
    h=dateheure(date0,heure0)-h0; printf("Processing time= %.6lf sec\n",h);
    } // loop While (!feof(pf1))
    // copy sum_radon to CPU
    cutilSafeCall(cudaMemcpy( radon_cum, sum_radon, 128*sizeof(float) , cudaMemcpyDeviceToHost ));
    float theta[128];
    float pas=180.00/128, t1, t2;
    float max=(float)(pas*128);
    theta[0]=0;
    for (i=1; i<128;i++){theta[i]=theta[i-1]+pas;}
    float S=0.0, pos=0.0; // S: Sensitivity; pos: theta at max
    calc_sens(theta,radon_cum, 128, &S,&pos);
    dm=tanf((90.0-pos)*3.14159265/180.0)*810.586*jnbbinav/64;
    sprintf(NOM_FICHIER, "GPDet-radon_cum_%08d_%03d-%04d-%03d.dat",id,jnbbinav,(int)dm,(int)S);
    fp = fopen(NOM_FICHIER, "w");
    fprintf(fp, "# Scan = %d Theta = %f° DM= %f SNR = %f\n",id,pos,dm,S);
    for (i = 0; i < 128 ; i++)
    {
        fprintf(fp, "%f \n",radon_cum[i]);
    }
    fclose(fp);
    } // if((pf1=fopen(st2,"r")) != NULL)

    fclose(pf1); printf("%s> Closing file %s\n",argv[0],filename);
    return(0);
    //Destroy CUFFT context
    cufftSafeCall(cufftDestroy(plan));
    free(data); free(c_fft_data); free(c_abs_fft_data); free(data_padded);
    free(c_radon); free(h_index);
    cutilSafeCall(cudaFree(d_data)); cutilSafeCall(cudaFree(d_fft_data));
    cutilSafeCall(cudaFree(d_abs_fft_data));
    cutilSafeCall(cudaFree(d_radon)); cutilSafeCall(cudaFree(sum_radon));
} // End of main

/*****
/*****      radon : Kernel performs the Radon transform      *****/
/*****
/*****

__global__ void radon( float* img_In, float* radonResults, int* index, int N, int M )
{
    // Declare rows of the matrix to be in shared mamory for speed
    __shared__ float sum[threads_per_block];

    // Calculate which element this thread reads from memory
    int idx1 = M * blockIdx.x + threadIdx.x;
    int idx = index[idx1];
    sum[threadIdx.x] = img_In[idx];
    __syncthreads();

    int nTotalThreads = blockDim.x; // Total number of active threads
    while (nTotalThreads > 1)
    {
        int halfPoint = (nTotalThreads >> 1); // divide by two
        // Only the first half of threads will be active.
        if (threadIdx.x < halfPoint)
            sum[threadIdx.x] += sum[threadIdx.x + halfPoint];
        __syncthreads();

        nTotalThreads = (nTotalThreads >> 1); // divide by two.
    }
    // At this time, each thread(0) has a sum of a row
    // It's time for each thread(0) to write it's final result.
    if (threadIdx.x == 0)
        radonResults[blockIdx.x]=sum[0];
}

/*****
/*****      cu_fftshift : shift the cufft quarter 1 to quarter 4 and quarter 2 to quarter 3 *****/
/*****

__global__ void cu_fftshift( float* img_In, float* img_Out, int* idx, int N, int M)
{
    //__shared__
    int idx1 = M * blockIdx.x + threadIdx.x;
    int id = idx[idx1];
    // blank 1rst row and 1rst column
    img_In[threadIdx.x] = 0;
    img_In[threadIdx.x*M] = 0;
    // shift
    img_Out[idx1] = img_In[id];
}

/*****
/*****      sqr_abs : Kernel calculates the square absolut of a complex value *****/
/*****

```

```

/*****
global void sqr_abs( cufftComplex * img_In, float* Result ,int M )
{
    // Calculate which element this thread reads from memory
    int idx = M * blockIdx.x + threadIdx.x;

    Result[idx] = (img_In[idx].x*img_In[idx].x)+(img_In[idx].y*img_In[idx].y);
}

/*****
global void cum_sum( float * vect, float* result ,int M )
{
    // Calculate which element this thread reads from memory
    int idx = M * blockIdx.x + threadIdx.x;

    result[idx] += vect[idx];
}

/*****
global void calc_index : calculates the vector of indexes used by Radon *****/
void calc_index( int M, int N, int *index )
{
    float pi = 3.14159265;
    unsigned int k, l=0;
    float pas=180.00/M, i, j;
    float max=(float)(pas*M);
    for (float theta=0; theta<max; theta+=pas)
    {
        if (theta == 135)
        {
            for (k = 0; k < N; k++)
            {
                j = k;
                i = N-1-j;
                // if (pow(j+(N+2)/2,2)+pow(i-(N+2)/2,2) < pow((N+2)/2,2))
                index[l*N+k] = floor(j+0.5) + M * floor(i+0.5);
            }
        }
        else if (theta == 45)
        {
            for (k = 0; k < N; k++)
            {
                j = k;
                i = j;
                index[l*N+k] = floor(j+0.5) + M * floor(i+0.5);
            }
        }
        else if (abs(theta-90) > 45) //probl me n'importe quelle theta m me somme
        {
            for (k = 0; k < N; k++)
            {
                j = k;
                i = tan (theta*pi/180)*(j-N/2) + M/2;
                // if (pow(j+(N+2)/2,2)+pow(i-(N+2)/2,2) < pow((N+2)/2,2))
                index[l*N+k] = floor(j+0.5) + M * floor(i+0.5);
            }
        }
        else
        {
            for (k = 0; k < N; k++)
            {
                i = k;
                j = tan ((90-theta)*pi/180)*(i-N/2) + M/2;
                // if (pow(j+(N+2)/2,2)+pow(i-(N+2)/2,2) < pow((N+2)/2,2))
                index[l*N+k] = floor(j+0.5) + M * floor(i+0.5);
            }
        }
        l++;
    }
}

/*****
global void shift_index : calculates the index *****/
void shift_index( int M, int N, int *idx)
{
    int *idx1, *idx2, *idx3, *idx4;
    int l;
    idx1 = (int*)malloc(sizeof(int)*M*N/4);
    idx2 = (int*)malloc(sizeof(int)*M*N/4);
    idx3 = (int*)malloc(sizeof(int)*M*N/4);
    idx4 = (int*)malloc(sizeof(int)*M*N/4);
    l=0;

```

```

for ( int i=0; i<N/2; i++) for ( int j=0; j<M/2; j++) {idx1[l] = j+N*i; l++;}
l=0;
for ( int i=0; i<N/2; i++) for ( int j=M/2; j<M; j++) {idx2[l] = j+N*i; l++;}
l=0;
for ( int i=N/2; i<N; i++) for (int j=0; j<M/2; j++) {idx3[l] = j+N*i; l++;}
l=0;
for ( int i=N/2; i<N; i++) for (int j=M/2; j<M; j++) {idx4[l] = j+N*i; l++;}

l=0;
for ( int i=0; i<N/2; i++) for ( int j=0; j<M/2; j++) {idx[j+N*i]=idx4[l]; l++;}
l=0;
for ( int i=0; i<N/2; i++) for ( int j=M/2; j<M; j++) {idx[j+N*i]=idx3[l]; l++;}
l=0;
for ( int i=N/2; i<N; i++) for (int j=0; j<M/2; j++) {idx[j+N*i]=idx2[l]; l++;}
l=0;
for ( int i=N/2; i<N; i++) for (int j=M/2; j<M; j++) {idx[j+N*i]=idx1[l]; l++;}

}

```


Bibliographie

- [Sup, 1999] (1999). Supplement to iee standard for information technology - telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements. part 11 : Wireless lan medium access control (mac) and physical layer (phy) specifications : High-speed physical layer in the 5 ghz band. *IEEE Std 802.11a-1999*.
- [Agee et al., 1990] Agee, B., Schell, S., and Gardner, W. (1990). Spectral self-coherence restoral : A new approach to blind adaptive signal extraction using antenna arrays. *Proceedings of the IEEE*, 78(4) :753–767.
- [Barnbaum and Bradley, 1998] Barnbaum, C. and Bradley, R. F. (1998). A new approach to interference excision in radio astronomy : real-time adaptive cancellation. *Astron. J.*, 115 :2598–2614.
- [Becker, 2009] Becker, W. (2009). *Neutron Stars and Pulsars*. Springer.
- [Bölcskei, 2001] Bölcskei, H. (2001). Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems. *IEEE Trans. on Communications*, 49(6) :988–998.
- [Boonsta, 2004] Boonsta, A. J. (2004). *Radio frequency interferences mitigation in radio astronomy*. PhD thesis, University of Delft, the Netherlands.
- [Boonstra and Weber, 2009] Boonstra, A. J. and Weber, R. (2009). Rfi mitigation methods inventory, skads european project. <http://webmail.jb.man.ac.uk/skadswiki/DS4Deliverables>.
- [Boonstra et al., 2010] Boonstra, A. J., Weber, R., and Colom, P. (2010). Rfi mitigation strategies for phased-array ska concepts.
- [Bretteil, 2005] Bretteil, S. (2005). *Estimation spectrale en temps réel de sources radioastronomiques en présence de brouilleurs cyclostationnaires*. PhD thesis, Université d’Orléans.
- [Cabric et al., 2004] Cabric, D., Mishra, S., and Brodersen, R. (2004). Implementation issues in spectrum sensing for cognitive radios. In *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on*, volume 1, pages 772 – 776.
- [Challapali et al., 2004] Challapali, K., Mangold, S., and Zhong, Z. (2004). Spectrum agile radio : Detecting spectrum opportunities. In *International Symposium on Advanced Radio Technologies March 2 - 4*, Boulder, Colorado, USA.
- [Ciblat and Serpendin, 2003] Ciblat, P. and Serpendin, E. (2003). A blind frequency offset estimator for OFDM/OQAM systems. In *Proc. of the IEEE SPAWC*, pages 595–599.
- [Deans, 1983] Deans, S. R. (1983). *The Radon Transform and Some of Its Applications*. Krieger Publishing Company, 1 edition.
- [Demorest, 2007] Demorest, P. B. (2007). *Measuring the Gravitational Wave Background using Precision Pulsar Timing*. PhD thesis, University of California, Berkeley.
- [Desvignes, 2009] Desvignes, G. (2009). *L’observation des pulsars au Radiotélescope de Nançay. Application à la recherche de nouveaux objets, l’étude des systèmes binaires relativistes et à la détection d’un font d’ondes gravitationnelles*. PhD thesis, Université d’Orléans.
- [Dumez-Viou, 2007] Dumez-Viou, C. (2007). *Restauration de sources radioastronomiques en milieu radioélectrique hostile : Implantation de détecteurs temps réel sur des spectres dynamiques*. PhD thesis, Université d’Orléans.
- [Eatough et al., 2009] Eatough, R. P., Keane, E. F., and Lyne, A. G. (2009). An interference removal technique for radio pulsar searches. *Monthly Notices of the Royal Astronomical Society*, 395(1) :410–415.
- [Eilek et al., 2004] Eilek, J. A., Hankins, T. H., and Jessner, A. (2004). Pulsar physics at low frequencies. *arXiv :astro-ph/0412278*.

- [Ellingson et al., 2001] Ellingson, J., and Bunton, S. W., and Bell, J. F. (2001). Removal of the glonass c/a signal from oh spectral line observations using a parametric modelling technique. *Astron. J. Supp*, 135(1) :87–93.
- [Feliachi, 2010] Feliachi, R. (2010). *a mettre*. PhD thesis, Université d’Orléans.
- [Ferréol et al., 2004] Ferréol, A., Chevalier, P., and Albera, L. (2004). Second-order blind separation of first and second-order cyclostationary sources - application to AM, FSK, CPSK and deterministic sources. *IEEE Trans. on Signal Processing*, 52(4) :845–861.
- [Fisher, 1936] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7 :179–188.
- [Fisher, 1938] Fisher, R. A. (1938). The statistical utilization of multiple measurements. *Annals of Eugenics*, 8 :376–386.
- [Foster et al., 1997] Foster, R. S., Ray, P. S., Cadwell, B. J., Backer, D. C., Lestrade, J. F., Cognard, I., Martin, J. M., and Maitia, V. (1997). A fast sampled high frequency pulsar survey of the galactic plane. In *Bulletin of the American Astronomical Society*, volume 29, page 1392.
- [Fridman, 2001] Fridman, P. A. (2001). Rfi excision using a higher order statistics analysis of the power spectrum. *Astron. Astrophys*, 368 :369–376.
- [Fridman, 2008] Fridman, P. A. (2008). Statistically stable estimates of variance in radio-astronomy observations as tools for radio-frequency interference mitigation. *The Astronomical Journal*, 135 :1810–1824.
- [Ganesan and Li, 2005] Ganesan, G. and Li, Y. (2005). Agility improvement through cooperative diversity in cognitive radio. In *Global Telecommunications Conference, 2005. GLOBECOM ’05. IEEE*, volume 5, page 2509.
- [Gardner, 1980] Gardner, W. A. (1980). A unifying view of second-order measures of quality for signal classification. *IEEE Transactions on Communications*, 28 :807–816.
- [Gardner, 1985] Gardner, W. A. (1985). *Introduction to random processes with applications to signals and systems*. Macmillan Publishing Company, New York, USA.
- [Gardner, 1987] Gardner, W. A. (1987). *Statistical spectral analysis : A nonprobabilistic theory*. Englewood Cliffs, Prentice-Hall, New Jersey, USA.
- [Gardner, 1991] Gardner, W. A. (1991). Exploitation of spectral redundancy in cyclostationary signals. *IEEE Signal Processing Magazine*, 8(2) :14–36.
- [Gardner et al., 2006] Gardner, W. A., Napolitano, A., and Paura, L. (2006). Cyclostationarity : Half a century of research. *Signal Processing*, 86(4) :639–697.
- [Gary et al., 2010] Gary, D. E., Liu, Z., and Nita, G. M. (2010). A wideband spectrometer with RFI detection. *Publications of the Astronomical Society of the Pacific*, 122 :560–572.
- [Geirhofer et al., 2006a] Geirhofer, S., Tong, L., and Sadler, B. (2006a). A measurement-based model for dynamic spectrum access in wlan channels. In *Military Communications Conference, 2006. MILCOM 2006. IEEE*, pages 1–7.
- [Geirhofer et al., 2006b] Geirhofer, S., Tong, L., and Sadler, B. M. (2006b). Dynamic spectrum access in WLAN channels : empirical model and its stochastic analysis. *Proceedings of the first international workshop on Technology and policy for accessing spectrum*, pages 14–es.
- [Ghozzi et al., 2006] Ghozzi, M., Dohler, M., Marx, F., and Palicot, J. (2006). Cognitive radio : methods for the detection of free bands. *Comptes Rendus Physique*, 7(7) :794–804.
- [Hansen and Toft, 1996] Hansen, K. V. and Toft, P. A. (1996). Fast curve estimation using preconditioned generalized radon transform. *IEEE Transactions on Image Processing : A Publication of the IEEE Signal Processing Society*, 5(12) :1651–1661.
- [Hellbourg et al., 2012] Hellbourg, G., Weber, R., Capdessus, C., and Boonstra, A.-J. (2012). Cyclostationary approaches for spatial RFI mitigation in radio astronomy. *Comptes Rendus Physique*, 13(1) :71–79.
- [Hewish et al., 1968] Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F., and Collins, R. A. (1968). Observation of a rapidly pulsating radio source. *Nature*, 217(5130) :709–713.
- [Leshem and van der Veen, 1999] Leshem, A. and van der Veen, A. J. (1999). *Introduction to interference mitigation techniques in radio astronomy*. ASTRON, the Netherlands.
- [Lorimer and Kramer, 2005] Lorimer, D. and Kramer, M. (2005). *Handbook of pulsar astronomy*. Astronomy and Astrophysics Library. Cambridge University Press, Cambridge, GB.

- [Lyne and Graham-Smith, 2006] Lyne, A. G. and Graham-Smith, F. (2006). *Pulsar astronomy*. Cambridge University Press.
- [McLaughlin et al., 2006] McLaughlin, M. A., Lyne, A. G., Lorimer, D. R., Kramer, M., Faulkner, A. J., Manchester, R. N., Cordes, J. M., Camilo, F., Possenti, A., Stairs, I. H., Hobbs, G., D’Amico, N., Burgay, M., and O’Brien, J. T. (2006). Transient radio bursts from rotating neutron stars. *Nature*, 439(7078) :817–820.
- [Napolitano and Spooner, 2001] Napolitano, A. and Spooner, C. M. (2001). Cyclic spectral analysis of continuous-phase modulated signals. *IEEE Trans. on Signal Processing*, 49(1) :30–44.
- [Nice, 1999] Nice, D. J. (1999). Radio pulses along the galactic plane. *The Astrophysical Journal*, 513(2) :927–932.
- [Nice and Taylor, 1995] Nice, D. J. and Taylor, J. H. (1995). Psr j2019+2425 and psr j2322+2057 and the proper motions of millisecond pulsars. *The Astrophysical Journal*, 441 :429–435.
- [Nita and Gary, 2010a] Nita, G. M. and Gary, D. E. (2010a). The generalized spectral kurtosis estimator. *Monthly Notices of the Royal Astronomical Society : Letters*, 406 :L60–L64.
- [Nita and Gary, 2010b] Nita, G. M. and Gary, D. E. (2010b). Statistics of the spectral kurtosis estimator. *Publications of the Astronomical Society of the Pacific*, 122 :595–607.
- [Nita et al., 2007] Nita, G. M., Gary, D. E., Liu, Z., Hurford, G. J., and White, S. M. (2007). Radio frequency interference excision using spectral-domain statistics. *Publications of the Astronomical Society of the Pacific*, 119 :805–827.
- [NVIDIA, 2008] NVIDIA (2008). *NVIDIA CUDA Compute Unified Device Architecture*.
- [Offringa et al., 2010] Offringa, A. R., de Bruyn, A. G., Biehl, M., Zaroubi, S., Bernardi, G., and Pandey, V. N. (2010). Post-correlation radio frequency interference classification methods. *astro-ph/1002.1957*.
- [Palicot, 2010] Palicot, J. (2010). *De la radio logicielle à la radio intelligente*. Hermes Sciences, 1 edition.
- [Poulsen et al., 2005] Poulsen, A. J., Jeffs, B. D., and Fisher, J. (2005). Programmable real-time cancellation of glonass interference with the green bank telescope. *Astron. J*, 130 :2916–2927.
- [Ray et al., 1999] Ray, P., Cadwell, B., Lazio, T., Foster, R., Backer, D., Cognard, I., and Lestrade, J. (1999). J0248+60 and j2240+58 : Two new pulsars from a high-frequency survey of the galactic plane. In *American Astronomical Society Meeting Abstracts #194*, volume 31 of *Bulletin of the American Astronomical Society*, page 903.
- [Raza et al., 2002a] Raza, J., Boonstra, A. J., and van der Veen, A. J. (2002a). Spatial filtering of RF interference in radio astronomy. *IEEE Signal Processing Letters*, 9(2).
- [Raza et al., 2002b] Raza, J., Boonstra, J., and van der Van, A. J. (2002b). Spatial filtering of rf interference in radio astronomy. *IEEE Signal Processing Letters*, 9(2).
- [Rickett, 1975] Rickett, B. (1975). Amplitude-modulated noise - an empirical model for the radio radiation received from pulsars. *apj*, 197 :185–191.
- [Sahai et al., 2006] Sahai, A., Tandra, R., Mishra, S. M., and Hoven, N. (2006). Fundamental design tradeoffs in cognitive radio systems. In *Proceedings of the first international workshop on Technology and policy for accessing spectrum*, TAPAS ’06, New York, NY, USA. ACM.
- [Serpedin et al., 2005] Serpedin, E., Panduru, F., Sari, I., and Giannakis, G. B. (2005). Bibliography on cyclostationarity. *Signal Processing*, 85(12) :2233–2303.
- [Shankar et al., 2005] Shankar, S., Cordeiro, C., and Challapali, K. (2005). Spectrum agile radios : utilization and sensing architectures. In *Proceedings of IEEE Int. Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pages 160–169, Baltimore, Maryland, USA.
- [Sonnenschein and Fishman, 1992] Sonnenschein, A. and Fishman, P. (1992). Radiometric detection of spread-spectrum signals in noise of uncertain power. *IEEE Transactions on Aerospace and Electronic Systems*, 28(3) :654–660.
- [Staelin and Reifenstein, 1968] Staelin, D. H. and Reifenstein, E. C. (1968). Pulsating radio sources near the crab nebula. *Science*, 162(3861) :1481–1483.
- [Tandra and Sahai, 2005] Tandra, R. and Sahai, A. (2005). Fundamental limits on detection in low snr under noise uncertainty. In *Wireless Networks, Communications and Mobile Computing, 2005 International Conference on*, volume 1, pages 464 – 469 vol.1.

- [Tang, 2005] Tang, H. (2005). Some physical layer issues of wide-band cognitive radio systems. In *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, pages 151 –159.
- [Theureau et al., 2011] Theureau, G., Parent, D., Cognard, I., Desvignes, G., Smith, D., Casandjian, J., Cheung, C., Craig, H., Donato, D., Foster, R., Guillemot, L., Harding, A., Lestrade, J., Ray, P., Romani, R., Thompson, D., Tian, W., and Watters, K. (2011). Psrs j0248+6021 and j2240+5832 : young pulsars in the northern galactic plane. discovery, timing, and gamma-ray observations. *Astronomy and Astrophysics*, 525 :A94.
- [Toft, 1996a] Toft, P. (1996a). *The Radon Transform - Theory and Implementation*. PhD thesis, Technical University of Denmark.
- [Toft, 1996b] Toft, P. (1996b). Using the generalized radon transform for detection of curves in noisy images. volume 4, pages 2219 –2222 vol. 4.
- [Urkowitz, 1967] Urkowitz, H. (1967). Energy detection of unknown deterministic signals. *Proceedings of the IEEE*, 55(4) :523 – 531.
- [Weber et al., 2004] Weber, R., Colom, P., Kerdraon, A., and Lecacheux, A. (2004). Techniques d’observation en radioastronomie basse fréquence en présence d’émetteurs radioélectriques. *Bulletin du Bureau National de Métrologie*, 126 :67–72.
- [Weber et al., 2007] Weber, R., Zarka, P., Ryabov, V. B., Feliachi, R., Griessmeier, J., Denis, L., Kozhyn, R. V., Vinogradov, V. V., and Ravier, P. (2007). Data preprocessing for decametre wavelength exoplanet detection : an example of cyclostationary rfi detector. In *Proceedings of EUSIPCO’09 conference*, Poznan, Poland.
- [Yilmaz and Doherty, 1987] Yilmaz, O. and Doherty, S. M. (1987). *Seismic Data Processing*. Society of Exploration.
- [Zeng and Liang, 2007] Zeng, Y. and Liang, Y.-C. (2007). Covariance based signal detections for cognitive radio. In *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on*, pages 202 –207.

Table des figures

1.1	(gauche) Modèle de l'étoile à neutrons et de sa magnétosphère avec l'émission radio s'échappant au niveau de l'axe magnétique ([Lorimer and Kramer, 2005]). (droite) Environnement immédiat du pulsar du Crabe (PSR B0531+21), en optique (image Hubble, en rouge) et en X (image Chandra, en bleu).	12
1.2	Le radiotélescope décimétrique de Nançay	13
1.3	Dispersion d'une impulsion individuelle du pulsar PSR B0329+54 observée à Nançay en avril 2009 sur une bande totale de 128MHz centrée à 1398MHz.	13
1.4	L'architecture typique d'une instrumentation 'pulsar' : a) étape 1 : acquisition de la forme d'onde b) étape 2 : banc de filtres polyphasés (PFB) pour séparer la bande de fréquence d'entrée (habituellement des centaines de MHz) en petits morceaux (B/N ~quelques MHz) c) étape 3 : après transformée de Fourier sur quelques MHz, soit un filtre approprié est appliqué pour dédisperser avant de revenir en temps pour effectuer la mesure de temps d'arrivée, soit une transformée de Fourier inverse particulière produit des séries temporelles sur de très petits canaux (B/NM ~0.25MHz) qui seront écrites sur disque en vue d'une exploration ultérieure en DM et en périodicité pour rechercher de nouveaux pulsars.	15
1.5	Plan temps fréquence d'une impulsion individuelle modélisée dans le domaine fréquentiel par l'équation 1.17. L'axe de fréquence est normalisé entre 0 et 1, alors que l'axe du temps l'est entre 0 et le nombre de points N.	17
1.6	Plan temps fréquence d'une impulsion individuelle modélisée dans le domaine temporel par l'équation 1.20. L'axe de fréquence est normalisé entre 0 et 1, alors que l'axe du temps l'est entre 0 et le nombre de points N.	18
2.1	Modélisation d'un récepteur de radioastronomie. Une succession de bancs de filtres (BF_k) fournit différents pavages temps-fréquence (T-F), $X_{\Delta t, \Delta f}^k(t, f)$ où k est un indice identifiant le banc de filtre, f et t sont respectivement les index fréquentiels et temporels qui positionnent le dit pavé T-F dans le plan T-F, Δf et Δt sont respectivement les résolutions fréquentielles et temporelles du pavage. Sauf besoin spécifique, le pavé T-F $X_{\Delta t, \Delta f}^k(t, f)$ sera simplement noté $X(t, f)$. Pour un signal initial de bande B , si nous considérons un point de mesure en début de chaîne, la résolution fréquentielle du plan T-F sera $\Delta f = B$ et la résolution temporelle $\Delta t = 1/B$. Par contre, en fin de chaîne, la résolution fréquentielle sera de $\Delta f = B/M$ et la résolution temporelle $\Delta t = M/B$ où M est le nombre de canaux fréquentiels que délivre la chaîne de mesure du radiotélescope.	22
2.2	Exemple de plan temps-fréquence (T-F). L'objectif des détecteurs de RFI est d'identifier les pavés T-F, $X(t, f)$, contenant une RFI (disque rouge) . Pour un détecteur donné, le nombre de disques rouges détectés sur le nombre de disques rouges total s'appelle la probabilité de détection P_d et toute détection d'un pavé autre que celui contenant une RFI s'appelle une fausse alarme dont on peut également calculer la probabilité P_{fa} . Un très bon détecteur est celui qui maximise P_d tout en minimisant P_{fa} : l'idéal étant d'obtenir $P_d = 1$ et $P_{fa} = 0$	22

2.3	Détecteur de puissance radar : (a) Schéma expliquant l'algorithme DetRadar. (b) Spectre de la raie 21 cm de l'hydrogène neutre (HI) de la galaxie PGC 51094 observée au NRT avec (noir) et sans (rouge) élimination radar. La figure a été tirée de [Dumez-Viou, 2007] avec son autorisation. Les paramètres de détection étaient les suivants : Bande d'analyse $\Delta f = 14$ Mhz, $N_{fort} = 3$, $N_{faible} = 30$, $N_{min} = 25$. Les probabilités de fausse alarme pour les indicateurs d'impulsions fortes et faibles étaient respectivement fixées à $2.1 \cdot 10^{-3}$ et $1.2 \cdot 10^{-3}$ via les paramètres respectifs $C_{fort} = 4$ et $C_{faible} = 0.8125$. La puissance de référence $\mu_N(t)$ était établie sur une fenêtre glissante de $N = 14000$ échantillons.	25
2.4	La modulation BPSK et ses caractéristiques cyclostationnaires. (a) Représentation temporelle schématisée d'un signal BPSK. Les paramètres de ce signal BPSK sont ceux utilisés pour les figures c et d. (b) Cas cyclique $\ R_{x,x}^\alpha(0)\ $, d'une observation au radiotélescope de Westerbork (Hollande) contenant un signal BPSK dû au satellite GPS. On retrouve la signature d'un des codes du GPS à 1023 MHz. (c) Cas cyclique $\ R_{x,x}^\alpha(0)\ $: (haut) bruit blanc gaussien seul, (bas) bruit blanc gaussien seul + BPSK (d) Cas cyclique conjugué $\ R_{x,x}^\alpha(0)\ $: (haut) bruit blanc gaussien seul, (bas) bruit blanc gaussien seul + BPSK. α est la fréquence cyclique normalisée, c'est à dire que $\Delta f = 1$ et $\Delta t = 1$	30
2.5	Distribution de probabilité d'un bruit blanc de puissance unité additionné à un signal BPSK (cf. Equ. 2.8) de différentes valeurs de puissance P_{BPSK}	31
2.6	Exemple de plan T-F de $x(t) = n(t) + p(t)$ avec $R_{PB} = 10$ dB. Les axes des fréquences et du temps sont normalisés (i.e. $\Delta f = 1$ et $\Delta t = 1$). Les amplitudes sont en dB. L'impulsion a un profil gaussien de largeur 64 échantillons. L'impulsion est dispersée sur la totalité des $N = 16384$ échantillons générés.	34
2.7	Schéma de principe pour calculer le profil simplifié de la réponse du détecteur (PSRD). Considérons l'histogramme des valeurs fournies par un détecteur pour un scénario donné. Le profil simplifié de l'histogramme est donnée par la position de la moyenne et par la plage de valeurs obtenues hors des 5% valeurs extrêmes. Cette formulation permet de voir la symétrie et l'étalement de la distribution.	35
2.8	Evaluation de la robustesse à la présence d'une impulsion pulsar. Les courbes représentent les profils simplifiés de l'histogramme des détecteurs sous H_0 et H_1 pour un R_{PB} variable. L'impulsion a un profil gaussien de largeur 64 échantillons. Elle est dispersée sur la totalité des $N = 16384$ échantillons qui servent à la détection. (a) Résultat du détecteur SK. (b) Zoom de (a). (c) Résultat de CycloDetFc. (d) Résultat de CycloDetFcA. (e) CycloDetBr. (f) CycloDetBrA.	36
2.9	(a) spectre cyclique, $\ R_{x,x}^\alpha(0)\ $, sous H_1 . (b) spectre cyclique sous H_0 . (c) 100 valeurs de CycloDetFcA sous H_0 et H_1	37
2.10	(a) spectre cyclique, $\ R_{x,x}^\alpha(0)\ $, sous H_1 . (b) spectre cyclique sous H_0 . (c) 100 Valeurs de CycloDetBrA sous H_0 et H_1	37
2.11	Exemples de RFIs de type BPSK mélangés à une impulsion pulsar et un bruit blanc gaussien. Les paramètres communs à cette simulation sont $N = 16384$, $R_{PB} = 0$ dB, $R_{IB} = -13$ dB et $F_c = 0.3142$. Les figures représentent le plan T-F $S(f, t)$ du signal en 2D (figures de gauche) et en 3D (figures de droite) (a) Cas du BPSK bande étroite, $Br = 1/100$ (i.e. un bit dure 100 échantillons) (b) Cas du BPSK bande large, $Br = 1/5$ (i.e. un bit dure 5 échantillons). Ici l'étalement du spectre de la RFI fait qu'elle est à peine visible en bas du plan T-F sous la forme d'une bande entre 0.2142 et 0.4142.	38
2.12	Test de Fisher des détecteurs en fonction de R_{IB} . Les paramètres communs de la simulation sont $N = 16384$, $R_{PB} = 0$ dB, $F_c = 0.3142$. Le Monte-Carlo s'est fait sur 100 essais. La ligne horizontale donne la limite pour obtenir une détection parfaite du RFI. (a) Cas du BPSK bande large avec $Br = 1/5$. (b) Cas du BPSK bande étroite avec $Br = 1/100$	39
2.13	Test de Fisher des détecteurs en fonction de N . Les paramètres communs de la simulation sont $R_{PB} = 0$ dB, $R_{IB} = 0$ dB, $F_c = 0.3142$. Le Monte-Carlo s'est fait sur 100 essais. La ligne horizontale donne la limite pour obtenir une détection parfaite du RFI. (a) Cas du BPSK bande large avec $Br = 1/5$. (b) Cas du BPSK bande étroite avec $Br = 1/100$	40
2.14	Spectres Cycliques et spectres cycliques conjugués avant et après dédispersion du signal. Les figures du haut correspondent aux résultats après dédispersion pour le spectre cyclique (a) et le spectre cyclique conjugué (c). Les figures du bas correspondent aux résultats avant dédispersion pour le spectre cyclique (b) et le spectre cyclique conjugué (d).	40

2.15	Schéma du protocole de test. Chaque pavé sera affecté par la valeur du critère de Fisher du scénario considéré. La sortie du banc de filtre (BF) numéro k contiendra $N_c = 2^k$ canaux de résolution fréquentielle, $\Delta f_k = \frac{1}{2^k}$ avec $k = 0, \dots, 8$. La dernière colonne indexée "ref" représente le profil spectral de la RFI seule. Elle permet de vérifier visuellement l'adéquation de ce profil avec les résultats de détection dans les différents pavés.	43
2.16	Test de Fisher des détecteurs en forme de "carte" dans le cas du pulsar seul. $RBP = 10dB$. L'objectif est de tester la robustesse des détecteurs à la présence d'une impulsion pulsar forte. .	45
2.17	Test de Fisher du détecteur CycloDetFcA en forme de "carte" en présence d'une interférence de type BPSK bande étroite ($Br = 1/100$) de très faible intensité ($RIB = -20dB$). Les autres paramètres de simulation sont $Fc = 0.31416$, $RBP = -20 dB$. Le brouilleur n'est détecté que lorsqu'il y a adéquation entre les caractéristiques spectrales de la RFI et la résolution spectrale du banc de filtres.	46
2.18	Test de Fisher des détecteurs, cas de la RFI BPSK bande étroite : $Br = 1/100$ $Fc = 0.31416$ $INR = 10 dB$ $PNR = -10 dB$	47
2.19	Test de Fisher des détecteurs en forme de "map" en présence d'une interférence de type BPSK $Br = 1/37$ $Fc = 0.31416$ $RIB = 10 dB$ $RPB = -10 dB$	48
2.20	Test de Fisher des détecteurs, cas de la RFI BPSK bande large : $Br = 1/5$ $Fc = 0.31416$ $RIB = 10 dB$ $RPB = -10 dB$	49
2.21	le projet UNIBOARD. (a) : La carte Uniboard, développée par les instituts partenaires ASTRON et JIVE, est constituée de 8 composants programmables (FPGA) de la société Altera (Altera Stratix IV EP4SGX230KF40). Chaque FPGA possède 4 liaisons 10GbitE pour se connecter directement vers l'extérieur grâce à des interfaces à liaison optique ou cuivre, soit un débit total 160Gbit/s pour chaque colonne. Chaque FPGA dispose également de deux banques de mémoire DDR3 de 4Go. (b) Présentation des différentes fonctions implantées dans la carte dans le cadre de l'application "récepteur de signaux pulsar ". Cette application est réalisée en partenariat avec l'université de Manchester	52
2.22	LOFAR à la station de Nançay. (a) Les antennes basse fréquence (30-80 MHz) de LOFAR. (b) Les "tuiles" haute fréquence (110-250 MHz).	53
2.23	Exemples d'application des détecteurs CycloDetBRA et SK pour l'observation du pulsar $B0329+54$ avec LOFAR, 992 canaux fréquentiels, bande passante $BW = 6.0552MHz$ entre $[142.9688, 149.0240] MHz$, $\Delta f = 6.104KHz$ et $\Delta t = 163.84\mu s$. SNR représente le rapport de la valeur du maximum détecté sur l'écart-type du bruit.	54
2.24	Le radiotélescope UTR-2 (The Ukrainian T-shaped Radio telescope).	55
2.25	Application sur les données de type "banc de filtres" pour les observations du pulsar $PSR B0031-07$: (a) le plan T-F de 2s des données pulsar $PSR B0031-07$ enregistrées à la bande de fréquence $[32.999737, 16.501049] MHz$ et échantillonnées à $4.028KHz$, $M = 4096$ canaux et $N = 256$ intants. (b) le plan T-F nettoyé en utilisant Kurtosis. (c) Plan T-F nettoyé en utilisant CycloDet. (d) Plan T-F nettoyé par les deux détecteurs.	56
3.1	Plan de dédispersion pour les données NBPP montrant l'étalement temporel de l'impulsions (en ms) en fonction de la mesure de dispersion DM et en prenant en compte toutes les contributions (élargissement produit par l'échantillonnage temporel et par la largeur du canal de fréquence). La courbe en rouge donne le pas optimum du DM (en $pc.cm^{-3}$, et la courbe en vert donne le pas adéquat de l'échantillonnage dt en ms	61
3.2	Description de l'algorithme SIPSFAR	62
3.3	Test de transformée de Radon sur un bloc de données contenant une impulsion géante du pulsar du Crabe	62
3.4	Illustration de l'intérêt de la transformation de Radon : (a) une image de taille $m \times n$ où quatre lignes sont noyées dans un bruit. (b) la TR de l'image précédente, le premier maximum à $\theta = 0^\circ$ représente la droite horizontale. Celui à 90° représente la droite verticale. Le maximum à 60° représente la droite ayant la pente positive et le dernier à 120° représente celle ayant la pente négative.	63
3.5	Illustration de la Discrete Radon Transform (DRT) pour $\rho = 0$: (a) une image avec une droite passant par le centre ayant une inclinaison par rapport à l'horizontale égale à 135° à partir de la gauche de l'image. (b) la DRT($\rho = 0, \theta$) appliquée à l'image avec une détection à $\theta = 135^\circ$. . .	64

3.6	Comparaison entre les méthodes SIPSFAR et PRESTO ($N = 256$, $SNR = -5dB$). Deux impulsions de durée $T = 1$ et $T = 8$ échantillons dans le plan t-f sont simulées pour différentes valeurs du nombre d'intégrations M . (a) Exemple d'un plan t-f utilisé pour la simulation. (b) Définition graphique des paramètres utilisés pour le calcul de la sensibilité. (c) Sensibilités pour différentes valeurs de T et de M . Les lignes continues et discontinues correspondent aux valeurs théoriques. Les carrés et les cercles correspondent aux valeurs calculées pour SIPSFAR en simulation.	66
3.7	Comparaison entre SIPSFAR et PRESTO : (a) Distribution des 4452 plans T-F des données du Crabe. La sensibilité de la méthode classique en fonction de la position du maximum dans la phase du profil du pulsar. Les cercles noirs indiquent la position des détections. Les données des plans T-F correspondant sont libellés $H1$. Les petits cercles gris représentent les données sans pulse dedans, et les données des plans T-F correspondant sont donc libellés $H0$. La figure juste à droite, représente le profil intégré obtenu avec les données $H1$. (b) La comparaison de la sensibilité des deux méthodes testées sur les données $H1$. La ligne discontinue verte représente la relation théorique entre les deux (voir 3.2.3). La ligne horizontale représente la limite inférieure donnant une détection avec une probabilité de fausse alarme (Pfa) égale à zéro. (c) La sensibilité obtenue après intégration de profils individuels ayant une sensibilité supérieure à un seuil donné. (d) Comparaison entre un profil individuel de faible sensibilité et un profil intégré.	68
4.1	Représentation en coordonnées galactiques des 38971 positions observées (sur 47229 potentielles) dans le cadre du programme de recherche de nouveaux pulsars mené en collaboration avec R.S.Foster du NRL Washington à la fin des années 1990. Ce programme a conduit à la découverte de deux pulsars jeunes [Theureau et al., 2011].	72
4.2	Trois secondes de spectre dynamique d'une observation de J0248+6021, un binning sur 128 échantillons a été appliqué en même temps qu'un DM de 360 pc.cm^{-3} pour la série temporelle du bas. Sept impulsions successives sont clairement visibles.	72
4.3	Architecture CUDA de l'exécution du kernel 'RADON'	74
4.4	Evolution de l'angle de Radon en fonction de la mesure de dispersion DM et du facteur de binning imposé. Pour chaque facteur de binning, les courbes ont été limitées pour éviter le recouvrement et donnent les angles de Radon à potentiellement utiliser pour éviter les détections multiples.	75
4.5	Exemple de spectre dynamique obtenu sur une portion de l'observation emplie de parasites.	76
4.6	Détections d'événements impulsionnels par PRESTO montrant une saturation irrécupérable. La technique PRESTO montre là une limite importante avec 20680710 événements.	76
4.7	Détections d'événements impulsionnels par la méthode RADON dans l'observation très parasitée. Ici, seuls deux événements (un à presque 50 sec et l'autre à 115 sec pour un DM proche de 0) sont détectés et démontre que la méthode est relativement insensible aux parasites.	77
4.8	Superposition des impulsions détectées par PRESTO (en gris) et par RADON (en noir) dans une observation du pulsar du Crabe (B0531+21, scan=21198027). Les trois impulsions assez faibles détectées par PRESTO ne le sont pas par SIPSFAR. A la date des observations, le radiotélescope n'avait pas encore été rénové et présentait une sensibilité 2.2 fois plus faible expliquant probablement la non-détection par la méthode SIPSFAR, car celle-ci a de faibles performances sur les signaux les plus faibles.	78
4.9	Superposition des impulsions dispersées détectées par PRESTO (en gris avec des détections à différents DMs) et par SIPSFAR (en noir avec une barre horizontale indiquant la largeur de la fenêtre de la FFT). Trois secondes d'une observation du pulsar J0248+6021 (scan 29298000) montrent sept (ou huit ?) impulsions successives toutes détectées également par RADON sauf la première alors que les trois dernières sont détectées toutes ensembles. (haut) tous les événements PRESTO sont représentés (bas) un regroupement des événements PRESTO a été effectué. On pourra utilement comparer avec la figure 4.2 qui montre la même portion de données.	80
4.10	Profil Radon de deux événements impulsionnels détectés sur le pulsar J0248+6021 (scan=29298000) pour différent facteur de 'binning'. (haut) quatre événements détectés simultanément (n° 2, 3, 4 et 5 de la figure 4.2) avec une fenêtre large. (bas) un événement isolé (n° 2) détecté avec une fenêtre étroite.	81
4.11	Superposition des impulsions détectées par PRESTO (en gris) et par SIPSFAR (en noir) dans l'observation de PSR J0248+6021 (scan 15798046). Si SIPSFAR détecte moins d'événements comme prévu car la technique est moins sensible, elle semble également détecter moins d'événements parasites.	82

4.12	Visualisation des différences de DM et de temps pour les couples d'événements (SIPSFAR,PRESTO) les plus proches.	82
4.13	Histogrammes des différences en DM : (noir) toutes les différences en temps sont prises en compte, (gris) les différences temporelles inférieures à 0.125 et 0.5 secondes.	83
4.14	Histogrammes des différences en temps : (noir) toutes les différences en DM sont prises en compte, (gris) les différences en DM inférieures à 10 et 50pc.cm ⁻³	83
B.1	Réponses impulsionnelles $h(t)$ (haut) et r_{h,h^*}^α (bas). Les cercles rouges correspondent aux positions des fréquences cycliques liées au baud rate de la modulation, ici $l/8$ avec $l = 0, \dots, 8$. On remarque que dans le cas d'un filtre rectangulaire, avec $\tau = 0$ seule la fréquence $\alpha = 0$ donne une amplitude non nulle.	91
C.1	La position d'une ligne en coordonnées polaires	93

Liste des tableaux

1.1	Extrait du tableau d'allocation des fréquences, ITU, Mai 2004	19
2.1	Tableau comparatif des méthodes d'estimation au second ordre. A part le filtrage "zéro DM", ces techniques présentent toutes le risque de confondre une impulsion géante de pulsar avec une RFI impulsionnelle.	27
2.2	Fréquences cycliques et fréquences cycliques conjuguées de différentes modulations ($l \in \mathbb{Z}$). . . .	31
4.1	Liste des 9 évènements SIPSFAR de l'observation 29298000 correspondants à des impulsions apparaissant à des DMs voisins de celui de J0248+6021 sur les 50 détectés au total.	79
4.2	Liste des 30 événements PRESTO correspondants à des impulsions réelles apparaissant à des DMs entre 340 et 420 pc.cm ⁻³ sur les 149 détectés au total. Notons que les 9 événements n° 2 à 10 correspondent aux 9 événements SIPSFAR listés en Table 4.1	79
4.3	Associations d'évènements PRESTO - SIPSFAR pour l'observation 2929800 de PSR J0248+6021. Comme attendu à partir des sensibilités théoriques des deux méthodes, tous les évènements SIPSFAR sont associés à des événements PRESTO alors que seuls 9 événements PRESTO sont associés à des événements SIPSFAR.	79
4.4	Associations avec des événements PRESTO à partir des événements SIPSFAR.	84



Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

Comptes Rendus Physique

www.sciencedirect.com



The next generation radio telescopes / Les radiotélescopes du futur

Blind detection of giant pulses: GPU implementation

Détection aveugle de pulses géants : implantation sur GPU

Dalal Ait-Allal^a, Rodolphe Weber^{b,a,*}, Cédric Dumez-Viou^a, Ismael Cognard^c,
Gilles Theureau^{a,c}

^a Observatoire de Paris, station de radioastronomie de Nançay, 18330 Nançay, France

^b Laboratoire PRISME, université d'Orléans, site Galilée, 12, rue de Blois, 45067 Orléans cedex 2, France

^c Laboratoire de physique et chimie de l'environnement et de l'espace, université d'Orléans/CNRS, 3, avenue de la recherche scientifique, 45045 Orléans cedex, France

ARTICLE INFO

Article history:

Available online 11 December 2011

Keywords:

RFI mitigation
Pulsar
Transient detection
GPU

Mots-clés:

Élimination des RFI
Pulsar
Détection de signaux impulsionnels
GPU

ABSTRACT

Radio astronomical pulsar observations require specific instrumentation and dedicated signal processing to cope with the dispersion caused by the interstellar medium. Moreover, the quality of observations can be limited by radio frequency interference (RFI) generated by Telecommunications activity. This article presents the innovative pulsar instrumentation based on graphical processing units (GPU) which has been designed at the Nançay Radio Astronomical Observatory. In addition, for giant pulsar search, we propose a new approach which combines a hardware-efficient search method and some RFI mitigation capabilities. Although this approach is less sensitive than the classical approach, its advantage is that no a priori information on the pulsar parameters is required. The validation of a GPU implementation is under way.

© 2011 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

RÉSUMÉ

L'observation radio des pulsars nécessite une instrumentation spécifique et des procédures de traitement du signal dédiées qui corrigent les effets de la dispersion induite par le milieu interstellaire. En outre, la qualité des observations peut être notablement dégradée par la présence d'interférences radioélectriques (RFI) d'origine anthropique. Ce papier présente l'instrumentation mise en place pour l'observation des pulsars à la Station de Radioastronomie de Nançay. Notamment, nous détaillons une approche originale pour la détection automatique de pulses géants. Bien que moins sensible que l'approche classique consistant à balayer en temps différé l'espace des paramètres du pulsar potentiel, l'approche proposée se distingue par une efficacité d'implantation, une capacité de traitement en temps réel et une robustesse intrinsèque aux RFI. En outre, elle ne nécessite aucune connaissance préalable des paramètres du pulsar, ce qui autorise son utilisation pour la détection d'événements impulsionnels non répertoriés. Une implantation sur GPU est en cours de validation.

© 2011 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

* Corresponding author at: Laboratoire PRISME, université d'Orléans, site Galilée, 12, rue de Blois, 45067 Orléans cedex 2, France.

E-mail addresses: dalal.ait_allal@obs-nancay.fr (D. Ait-Allal), rodolphe.weber@univ-orleans.fr (R. Weber), cedric.dumez-viou@obs-nancay.fr (C. Dumez-Viou), icognard@cnrs-orleans.fr (I. Cognard), theureau@obs-nancay.fr (G. Theureau).

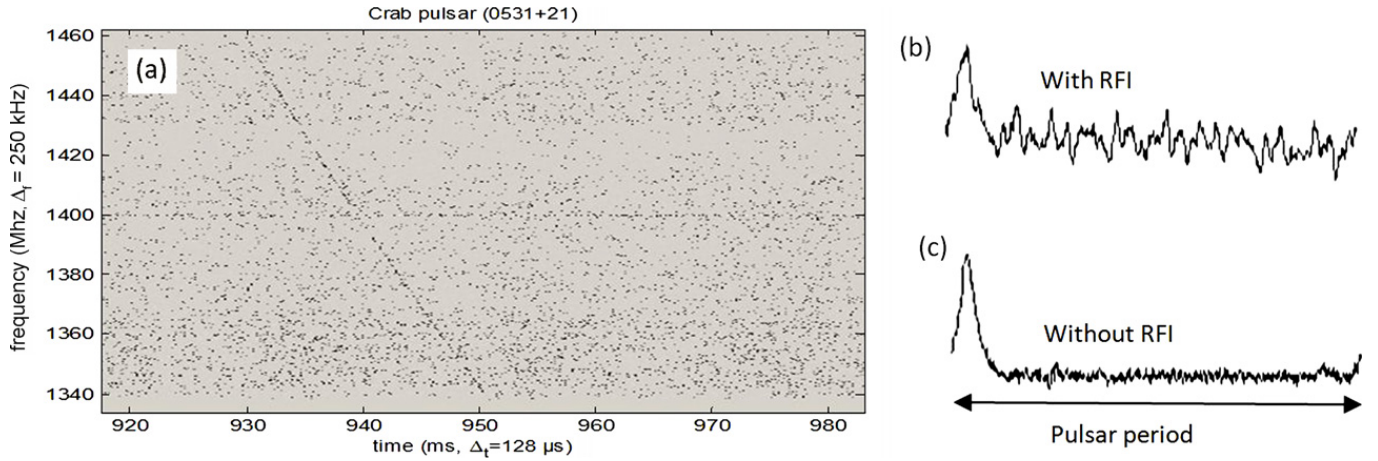


Fig. 1. Pulsar examples. (a) Time–frequency representation of one giant pulse of the Crab pulsar. (b) Time representation of the average power pulse profile (pulsar J0034-0721) after coherent de-dispersion but with RFI signals present in the observed band. (c) Time representation of the average power pulse profile of the pulsar J0034-0721 after coherent de-dispersion and with RFIs blanked out.

1. Introduction

Pulsars are rapidly rotating (up to 716 rotations per second) highly magnetized, neutron stars which produce radio beams which sweep the sky like a lighthouse. If the beam is oriented toward the Earth, it will produce periodic pulses which can be measured with radio telescopes and dedicated backends [1]. As they travel through the ionized interstellar medium (ISM), these pulses are progressively attenuated and spread over time (i.e. the higher pulsar frequencies will arrive earlier than the lower pulsar frequencies). Fig. 1(a) shows a raw time–frequency representation of one of these pulses. This phenomenon, named dispersion, makes these pulses barely detectable without further processing. Moreover, the quality of pulsar observations is also limited by radio frequency interferences (RFI) generated by various (and growing) telecommunications activities. Fig. 1(c) gives an example of the distortion induced by RFI on pulsar observation (to be compared with Fig. 1(b)). To cope with RFI, different RFI blanking techniques can be implemented. Examples of RFI mitigation procedures are given in [2,3].

In Section 2, the functional description of the coherent de-dispersion algorithm is explained and the innovative hardware architecture of the Nançay de-dispersion backend is detailed. When both the dispersion measure (DM, i.e. the integrated column density of free electrons between an observer and a pulsar) and the period of the pulsar are known, this architecture makes a precise timing of the corresponding pulsar possible.

The search for new pulsars is a difficult task since neither the DM nor the periodicity is known. The classical approach consists in recording all the channelized data and in incoherently de-dispersing them by an off-line search in the DM space. Finally, all the data are folded with different sets of periods, providing a list of pulsar candidates. This approach requires huge disk space and is quite time consuming.

In Section 3, we propose a new real time pulsar search method, requiring small disk space, fewer computational resources and providing a simple way to blank RFI. Its advantage is that no a priori information on the pulsar parameters is required.

As will be shown in Sections 4 and 5, the drawback is that this approach is less sensitive than the classical approach mentioned above. Consequently, the proposed approach is well adapted to high dynamic transient emissions such as rotating radio transients (RRATs, short and bright single pulses observed one at a time with the average time intervals between bursts ranging from a few minutes to several hours [4]) or giant pulses (very bright pulses occurring from time to time [5,6]).

In Section 6, the GPU implementation of this algorithm is described.

2. The coherent de-dispersion receiver

2.1. Coherent de-dispersion principle

ISM dispersion can be modeled as a frequency transfer function [1], $H(f)$:

$$H(f + f_0) = \exp\left(j \frac{2\pi DM f^2}{2.41 \times 10^{-10} (f + f_0) f_0^2}\right) \quad (1)$$

where f_0 is the center frequency of observed bandwidth Δf and $|f| < \Delta f/2$. Coherent de-dispersion consists in applying the inverse of this transfer function to the dispersed received signal. This de-dispersion is done in the complex frequency

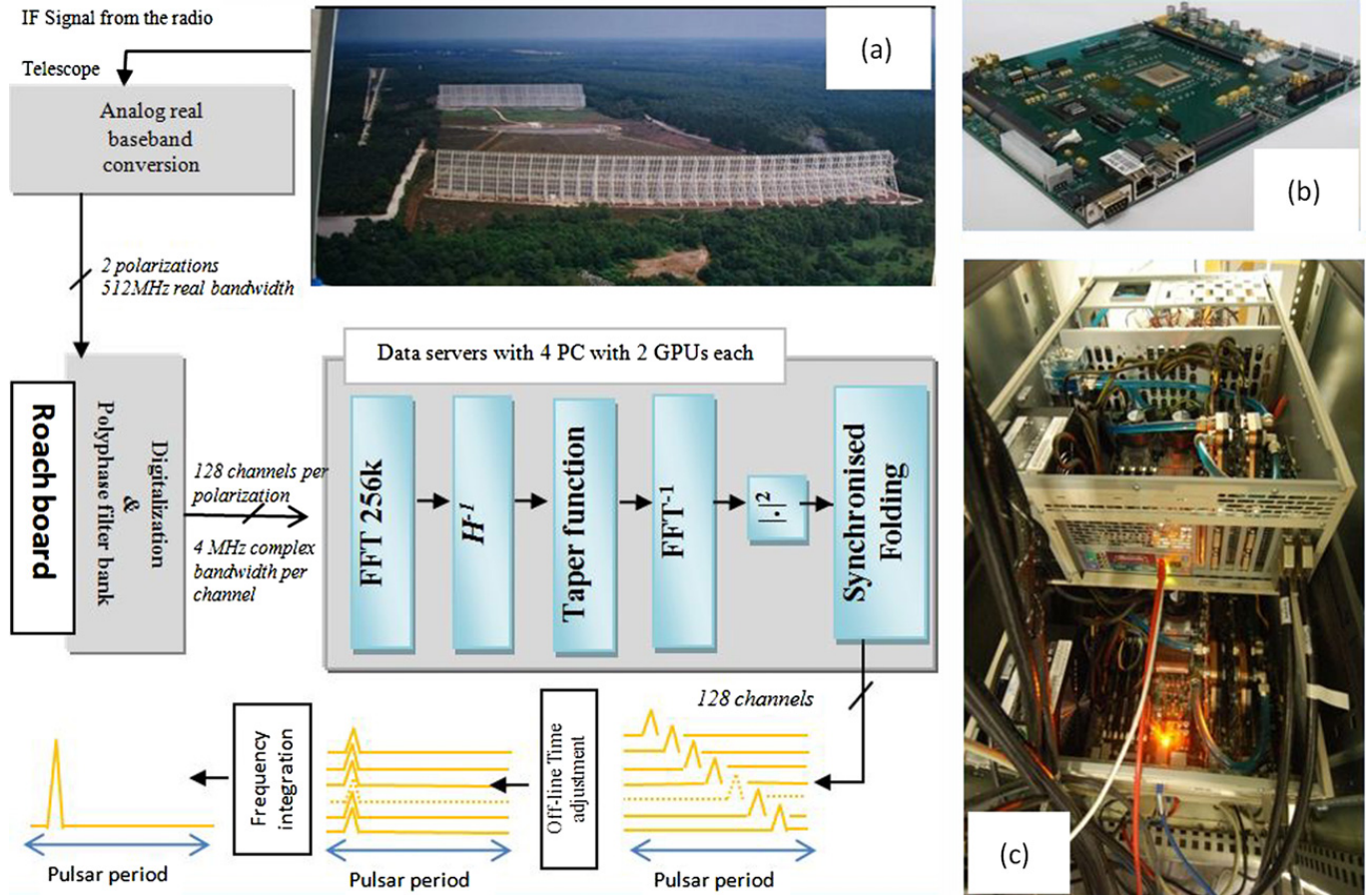


Fig. 2. Nançay coherent pulsar receiver. (a) Description of the instrumentation. (b) ROACH digital board for the waveform digitization and the polyphase filter bank. (c) “Data servers” with 4×2 GPUs which compute the coherent de-dispersion and the synchronized folding on each frequency channel.

domain through Fourier transforms. The operations are detailed inside the “Data server” box in Fig. 2(a). A taper function is used to avoid aliasing in the low-pass filtering.

2.2. Receiver hardware implementation

Fig. 2 shows the signal data-flow from the radio telescope to the final de-dispersed pulses. The different steps are described below:

- The Analog System: two orthogonal polarization signals (500 MHz bandwidth) from the radio telescope are downconverted for analog to digital conversion and channelizing.
- The ROACH Board (cf. Fig. 2(b) and [7]): this digital board was designed by CASPER (Center For Astronomy Signal Processing and Electronics Research). For this application, it contains two 8-bit analog to digital converters (ADCs) which are clocked at 1024 MSa/s. A logic programmable device (Virtex 5 SX95T from Xilinx) is used to perform a 128 channel polyphase filter bank (PFB). Each channel output bandwidth of complex data samples is 4 MHz.
- Data servers (cf. Fig. 2(c)): this is the innovative part of the design. Four PCs embedding two GPUs each (Graphic Processing Unit – NVIDIA GT280 240 parallel processors) are used for the final processing steps. Each PC receives a quarter of the bandwidth through 10 GbE (Gigabit Ethernet) links. The data are then read by the de-dispersion program which performs the following tasks:
 - Convert raw data from proprietary fixed point format to floating point.
 - Apply a Fourier transform of appropriate length (256k bins).
 - Multiply by the de-dispersion filter, $H(f)^{-1}$ and the taper function.
 - Inverse Fourier transform back to the time domain.
 - Detect the data to obtain power versus time, and crossmultiply polarization terms.
 - Folding at the pulsar period with a 60 s dump time.
- Channel integration to enhance the sensitivity.

This system is now fully operational at Nançay Observatory. In terms of sensitivity, it outperforms the previous system presented in [2] by a factor of 2.

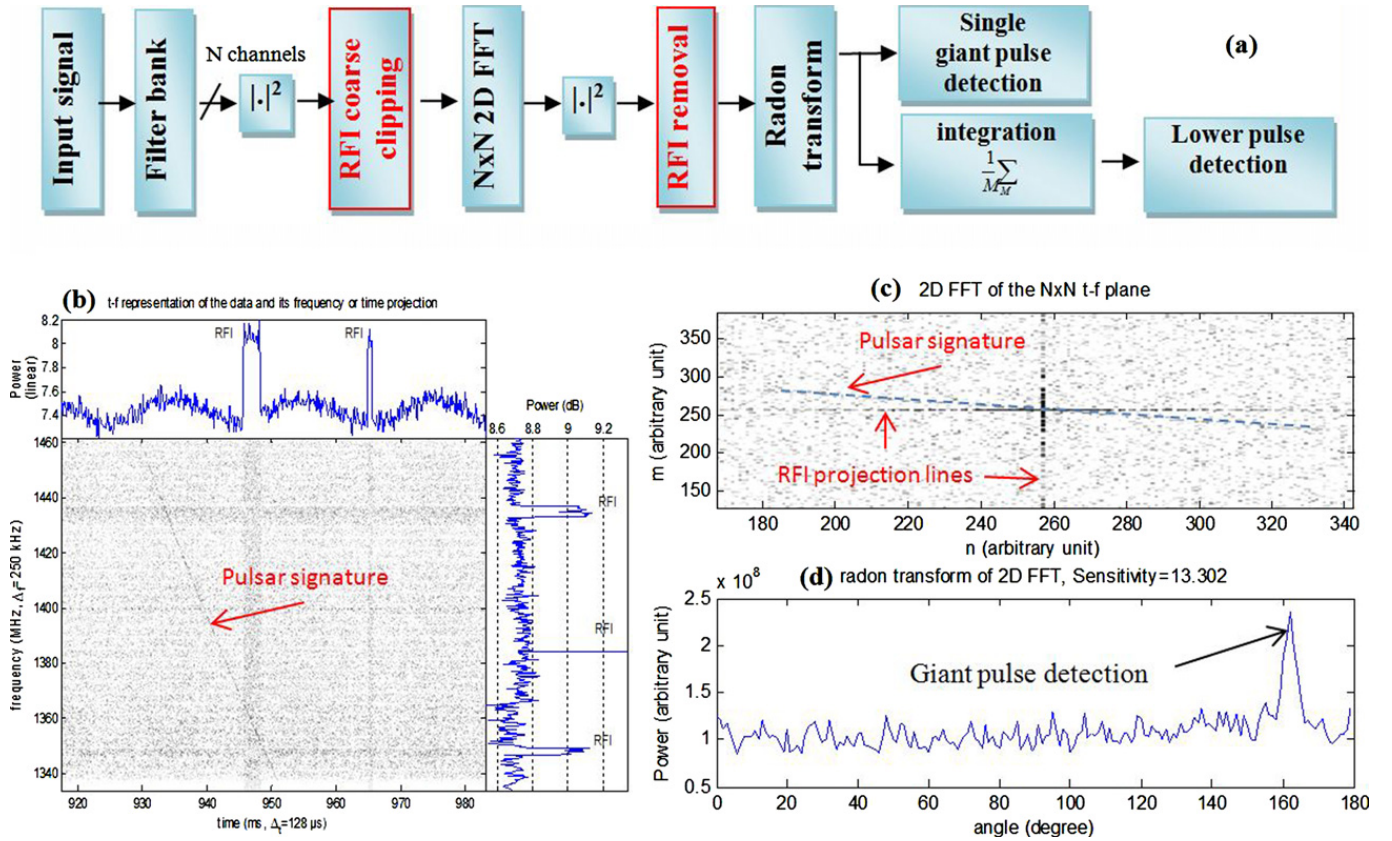


Fig. 3. (a) Algorithm description. (b) Time–frequency (t–f) representation ($N = 512$) of the signal. On the integrated power figures (top and right), the time and frequency RFI positions can be easily located. The three RFI spectral lines and the two RFI time bursts are simulated RFIs. The dispersed pulse is emitted by the Crab pulsar. (c) Example of the 2D-FFT output (zoomed in). All the RFIs are concentrated on the central cross. The oblique line is the signature of a dispersed pulse. Its angle depends on the pulsar DM. (d) Radon transform of the previous cleaned 2D-FFT plane (i.e. the central cross has been blanked). The peak indicates the presence of a pulse. The result is strictly similar if there is no RFI.

3. The giant pulse blind detector

In Fig. 3, the different steps of the algorithm are presented and an example with a real pulsar pulse is provided for illustration. In short:

- After analog to digital conversion, a time–frequency (t–f) power plane of the waveform is produced through a digital filter bank (see an example in Fig. 3(b)). This is a regular real-time functionality in all radio telescopes.
- To remove obvious RFI, t–f slots with high levels are truncated. A simple power detector with a coarse threshold can be implemented. If computational resources are available, a more sophisticated approach may be chosen [8,9].
- Real time 2D Fourier transforms (2D-FFT) are applied on successive $N \times N$ t–f images. An example is given in Fig. 3(c). With this approach, all the dispersed pulsar pulses are projected at the same location in the 2D-FFT image. This location depends only on the pulsar DM and is different from the one obtained for RFI. In other words, any impulsive (respectively continuous) RFI will be concentrated in the horizontal (resp. vertical) line centered in the 2D-FFT image. Thus, to remove RFI and to obtain a clean image, all that is required is to blank these vertical and horizontal lines which cross at the center of the 2D-FFT image.
- A Radon transform is applied on the cleaned 2D-FFT image. It consists in summing the image intensity along successive radial lines [10]. When this integration line corresponds to the line where the pulsar pulses have been projected, all the pulsar power is concentrated in one point. At other integration angles, only the noise contribution will be integrated. Fig. 3(d) illustrates this principle.

The advantage of this approach is that only a limited amount of data (the Radon profile of the $N \times N$ 2D-FFT image) is stored for further processing. Furthermore, it provides simple RFI mitigation capabilities. The successive Radon profiles can also be integrated over time to enhance the detection of pulses. In the following, we investigate the sensitivity of this 2D-FFT/Radon approach compared with the classical one (i.e. incoherent de-dispersion).

4. Sensitivity comparison

We consider the following signal model and algorithm parameters:

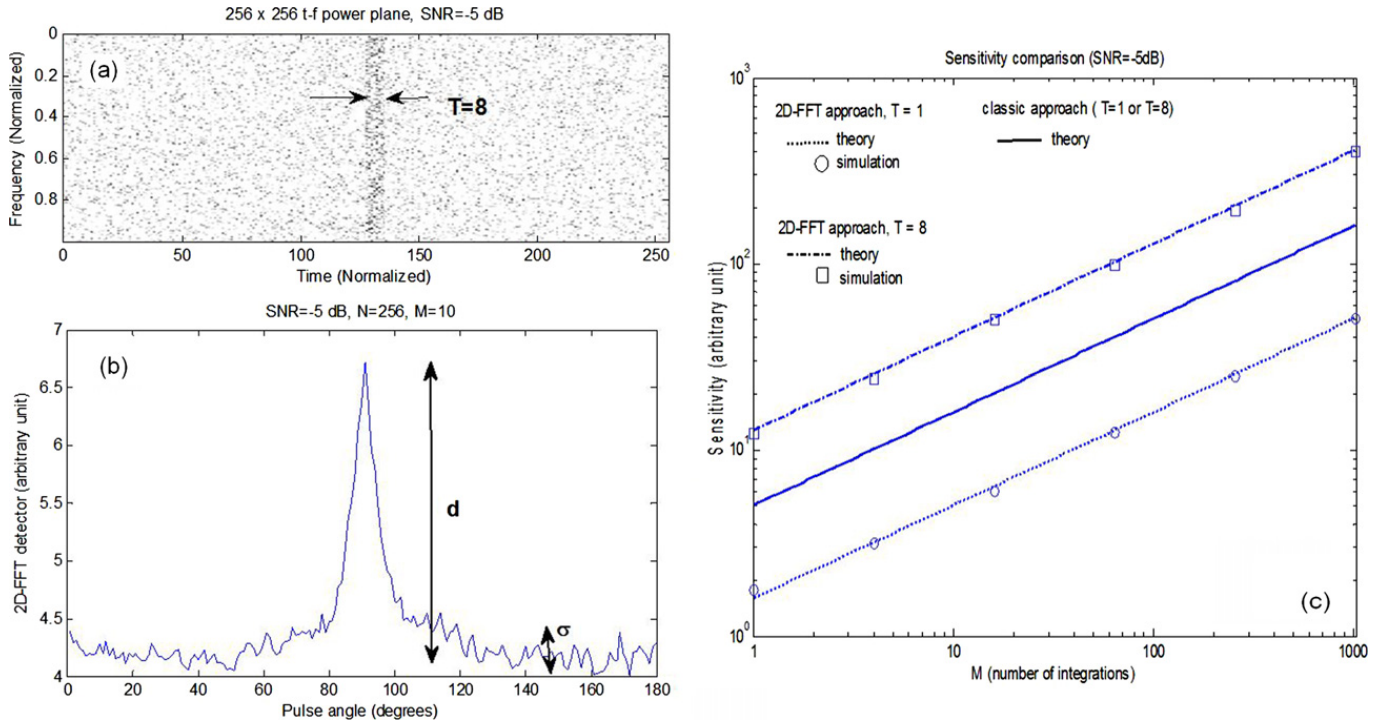


Fig. 4. Comparison of the 2D-FFT and the classical pulsar search procedures ($N = 256$, $\text{SNR} = -5$ dB). Two pulse lengths ($T = 1$ and $T = 8$ in samples) were simulated for different values of M . (a) Model of the t-f plane used for the simulations. (b) Graphical definition of the parameters used to compute the sensitivity. (c) Sensitivities for different values of T and M . The continuous and dashed curves correspond to theoretical values. Boxes and circles correspond to 2D-FFT sensitivity measurements through simulations.

- The background noise is a white zero-mean Gaussian noise with power σ_n^2 .
- The pulsar pulse is a white zero-mean Gaussian noise with power σ_p^2 . The pulse duration is L . The signal to noise ratio is defined by $\text{SNR} = \sigma_p^2 / \sigma_n^2$. For the theoretical sensitivity derivation, we assume that $\sigma_p^2 \ll \sigma_n^2$.
- There is no RFI.
- The t-f power plane size is $N \times N$. The time and frequency units are normalized.
- For the classical approach, we consider that the DM and the pulsar periodicity have been found by the ad-hoc iterative procedure. Thus, it is equivalent to assuming: (1) $\text{DM} = 0$ (i.e. vertical pulse); and (2) pulses always occur at the same position in the t-f power plane. These t-f power planes are integrated over frequency to provide individual pulse profiles. Finally, M pulse profiles are summed together. The sensitivity, S_{classic} , is defined by the average profile deviation due to the pulse over the noise only profile standard deviation.
- For the 2D-FFT approach, we set DM to zero as well. Indeed, with no loss of generality, we consider that zero-DM is just a DM among others. The central pixel of the 2D-FFT image is systematically blanked. The average of M 2D-FFT planes provides the final clean 2D-FFT image. Finally, the Radon transform is applied for different angles. The sensitivity, $S_{2\text{D-FFT}}$ is defined by the Radon transform deviation at angle 90° over the Radon transform standard deviation at other angles.

This model is illustrated in Fig. 4(a). A graphical definition of sensitivity is provided in Fig. 4(b). By considering the mean and the variance of this model at the successive algorithm stages, we have derived the following theoretical equations:

- for the classical case: $S_{\text{classic}} = \frac{\sigma_p^2}{\sigma_n^2 / \sqrt{NM}} = \text{SNR} \sqrt{NM}$;
- for the 2D-FFT case: $S_{2\text{D-FFT}} = \frac{T \sigma_p^4}{\sigma_n^4 / \sqrt{NM}} = T \cdot \text{SNR}^2 \sqrt{NM}$ where $T = L/N$ (i.e. T is the pulse width in the t-f power plane).

In Fig. 4(c), these sensitivities were computed by simulation for different parameters. Theoretical sensitivities are also plotted. From this figure, several remarks can be made:

- It is shown that there is good agreement between theory and simulations.
- The 2D-FFT sensitivity is proportional to SNR^2 . This makes this approach less sensitive to low SNR compared with the classical approach (see for example the case ($\text{SNR} = -5$ dB, $T = 1$)). This drawback is counterbalanced by the fact that the 2D-FFT approach is also sensitive to the pulse width, T . The wider the pulse is, the better the 2D-FFT detection is (see case ($\text{SNR} = -5$ dB, $T = 8$)).

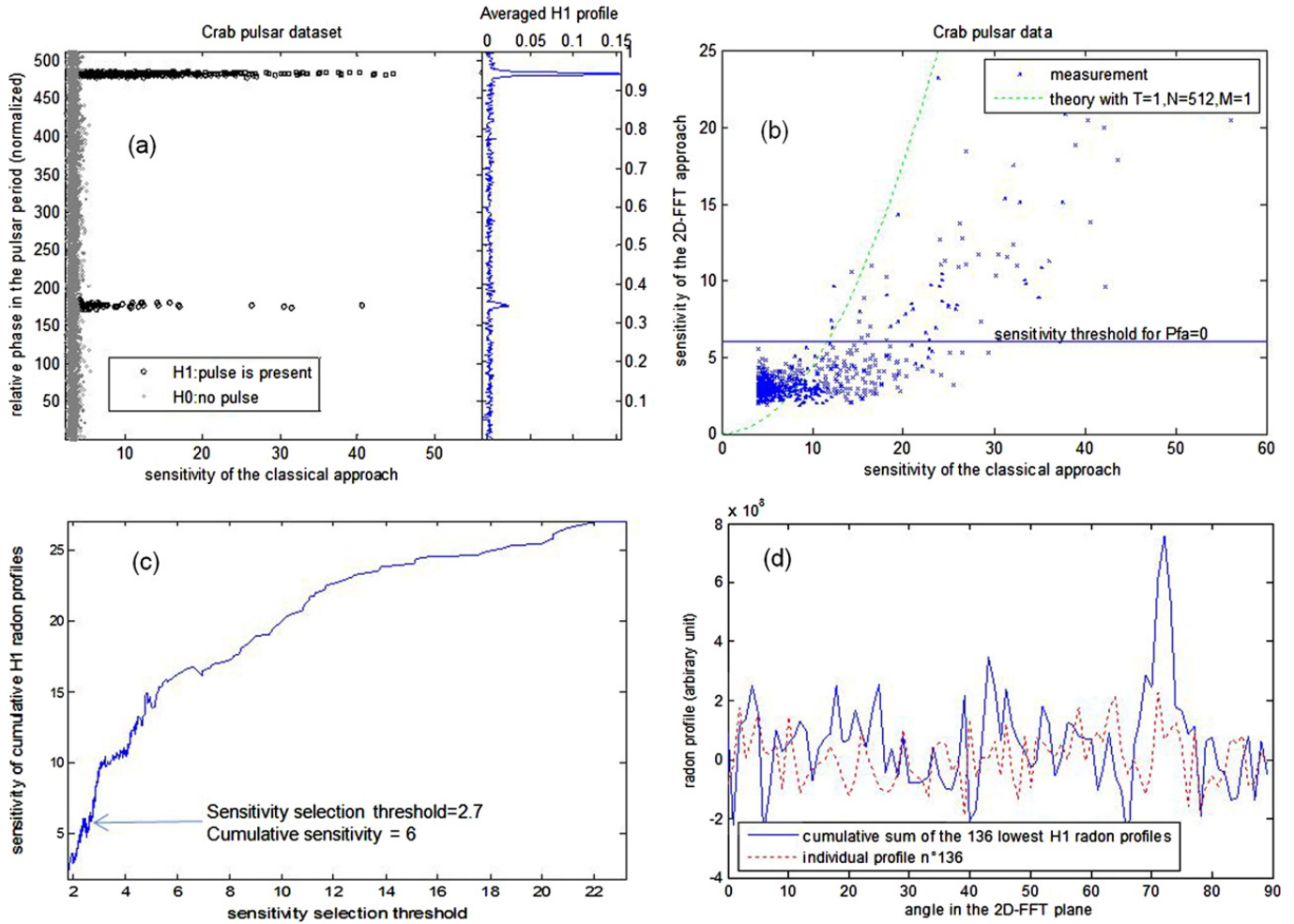


Fig. 5. (a) Distribution of the 4452 t-f planes from the Crab pulsar data set. Classical (incoherent de-dispersion) sensitivities obtained versus the corresponding maximum position in the de-dispersed pulse profile. Black circles indicate positive detections. The corresponding t-f planes are labeled H1 data set. Small gray circles represent the H0 data set with no pulse inside the t-f planes. The plot on the right gives the integrated pulse profile obtained with the H1 data set. (b) Sensitivity comparison between the classical approach and the 2D-FFT approach with the H1 data set. The dashed line represents the theoretical relation between these sensitivities (see Section 4). The horizontal line represents the positive detection limit providing a probability of false alarm (Pfa) equal to zero. (c) Sensitivity obtained after integration of all the individual Radon profiles with a sensitivity below a given threshold. (d) Comparison of individual and cumulative Radon profiles for the parameters indicated with an arrow in (c).

- These detection performances are achieved without requiring any prior knowledge about the pulsar (neither the DM nor the periodicity). This is not the case for the classical pulsar search procedure, where the derived performances are only achieved when exact DM and periodicity are known.

5. Tests on Crab pulsar data

The two detectors were applied to a Crab pulsar data set acquired with the Nançay Decimeter Radio Telescope. This data set corresponds to 4452 t-f planes generated with the size parameter $N = 512$. The $N = 512$ frequency channels cover the band from 1.3339 GHz to 1.4619 GHz. The temporal resolution is 128 μ s. Consequently, the total duration of the data set is 4 min 55 s. The pulsar parameters applied for the incoherent de-dispersion process (classical approach) are $DM = 56.65 \text{ pc cm}^{-3}$ and period = 33.63128 μ s.

The classical approach is used to identify t-f planes with at least one pulse inside. The detection criteria are “*sensitivity above 4*” and “*peak position at the expected pulse positions within the pulsar period*”. In Fig. 5(a), the black circles represent the 589 t-f planes fulfilling the previous detection thresholds. This set of t-f planes is called the H1 data set. The 3863 other t-f planes without pulse inside (small gray circles) are called the H0 data set. Fig. 5(a) on the right shows the averaged de-dispersed pulse profile obtained from the H1 data set.

In Fig. 5(b), the sensitivities measured in the H1 data set between the classical and the 2D-FFT approaches are compared. It can be seen that the 2D-FFT performance is appreciably lower than the expected theoretical one (see Section 4). One reason is that real data sets are affected by propagation issues such as scintillation. Another one is that real pulse profiles are not perfectly rectangular as was assumed in Section 4. However, the positive point is that with a sensitivity threshold around 6, the 2D-FFT detector is able to detect strong pulses with no false alarm. More precisely, 65 t-f planes provide

2D-FFT sensitivities above 6 (compared to the 365 classical detections obtained with the same sensitivity threshold). A first conclusion is that very strong pulsar pulses can be detected with no false alarm and without any prior knowledge of the pulsar DM. Furthermore, the angle corresponding to the detection peak in the Radon profile provides a first estimate of the unknown pulsar DM.

Figs. 5(c) and 5(d) show that detection with lower sensitivities can be achieved by integrating Radon profiles. Thus, Fig. 5(c) shows that by summing all the individual H1 Radon profiles with a sensitivity lower than a given sensitivity threshold, a new Radon profile can be generated with a better signal to noise ratio. For example, by integrating the 136 lowest individual H1 Radon profiles (i.e. sensitivity threshold = 2.7), a zero false alarm detection is obtained (i.e. new sensitivity = 6). For this example, one individual profile (dashed red line) and the cumulative one (continuous blue line) are compared in Fig. 5(d). Thus, the second conclusion is that the proposed approach can achieve positive detections of lower pulses by integrating the individual Radon profiles. Similarly, the classical detector can be enhanced by folding the de-dispersed profile. However, the knowledge of pulsar period is required in that case.

This study on real data confirms that the 2D-FFT approach is an interesting alternative for the blind detection of strong pulses in real time.

6. GPU implementation

Performance tests were conducted to prove the effectiveness of GPUs as an appropriate solution to data processing for wider band and in real time [11]. The new blind detection method is implemented on a machine equipped with a GeForce GTX 85, with 240 cores, 1 GB of global memory and 16 MB of shared memory. The operating system is Gentoo Linux with CUDA library 2.3. The CUDA library is very well suited to exploiting the parallel computing capabilities of the GPU. Indeed, when the GPU is programmed through CUDA C, which is an extension of the C language, it becomes a computing device capable of performing a large number of parallel operations; an algorithm can be executed several times on several independent data sets. In the new method explained in Section 3, two steps are required for its implementation: the two-dimensional Fourier transform (FFT^2) and the Radon transform. The FFT can be done with the CUFFT library which comes with CUDA C. In this section, we will therefore focus on the implementation of the Radon transform. In fact, the Radon transform is simply the sum of intensities in the image as a function of the radial lines through the center of the image. Call θ the angle of the radial with the horizontal. Due to the symmetry present in the output image of FFT^2 , and in order to minimize the computation time, the Radon transform is applied on only a quarter of the image between $\theta = 0^\circ$ and $\theta = 90^\circ$. Then, for each value of θ , the pixels index of the corresponding radial line is calculated in advance. The K indexes of the N angles θ are then stored in an index matrix ($N \times K$) and transferred once and for all in the shared memory of the GPU. This technique optimizes data transfer and improves processing speed. Finally, the sum of the intensities according to each row of the index matrix is calculated by the reduction algorithm, which is the sum of two by two elements at each stage, thus halving the number of calculations every time. The result is a vector containing the cumulative intensities for each value θ as shown in Fig. 3(d).

7. Conclusions

A new approach for pulsar search, combining a hardware-efficient search method and some RFI mitigation capabilities has been proposed. It is based on two-dimensional Fourier transform and Radon transform. While this algorithm is less sensitive than the classical one based on incoherent de-dispersion, positive detection can be achieved on strong pulses without knowing the DM of the corresponding pulsar. It could be an alternative for the next generation of radio telescopes such as the LOFAR radio telescope (www.lofar.org) or the Square Kilometre Array (SKA, www.skatelescope.org) where the huge amount of data to be processed by classical search procedures is a serious issue.

Acknowledgements

The authors would like to thank the European Commission Framework Program 7, Project PrepSKA (contract No. 212243) and the French funding agency ANR (contract ANR-09-BLAN-0225-04) for funding part of this work.

References

- [1] D. Lorimer, M. Kramer, Handbook of Pulsar Astronomy, Cambridge University Press, Cambridge, 2005, pp. 106–123 and 141–142.
- [2] D. Ait-Allal, C. Dumez-Viou, R. Weber, G. Desvignes, I. Cognard, G. Theureau, RFI mitigation at Nançay Observatory: Impulsive signal processing, in: S.A. Torchinsky, A. van Ardenne, T. van den Brink-Havinga, A. van Es, A.J. Faulkner (Eds.), Widefield Science and Technology for the SKA, ISBN 978-90-805434-5-4, 2010.
- [3] R.P. Eatough, E.F. Keane, A.G. Lyne, An interference removal technique for radio pulsar searches, Mon. Not. R. Astron. Soc. 395 (2009) 410.
- [4] M.A. McLaughlin, et al., Transient radio bursts from rotating neutron stars, Nature 439 (7078) (2006) 817–820.
- [5] D.H. Staelin, E.C. Reifstein, Pulsating radio sources near the Crab nebula, Science 162 (3861) (1968) 1481–1483.
- [6] I. Cognard, et al., Giant radio pulses from a millisecond pulsar, ApJL 457 (1996) L81.
- [7] <https://casper.berkeley.edu/wiki/ROACH>.
- [8] R. Weber, et al., Data preprocessing for decametre wavelength exoplanet detection: an example of cyclostationary RFI detector, in: EUSIPCO'2007, Poznan, Poland, August 2007.

- [9] G.M. Nita, et al., Radio frequency interference excision using spectral-domain statistics, *Publ. Astron. Soc. Pacific* 119 (857) (2007) 805–827.
- [10] S. Deans, *The Radon Transform and Some of Its Applications*, Krieger Publ. Co., 1983.
- [11] G. Desvignes, *L'observation des pulsars au Radiotélescope de Nançay. Applications à la recherche de nouveaux objets, à l'étude des systèmes binaires relativistes et à la détection d'un fond d'ondes gravitationnelles*, PhD thesis, University of Orléans, 2009.

Exemples de traitement en temps réel des interférences radioélectriques pour la radioastronomie

RODOLPHE WEBER^{1,2}

AVEC LA COLLABORATION DE CÉDRIC DUMEZ-VIOU¹, SAMAR CHANGUEL^{1,2} DALAL AIT-ALLAL¹

¹OBSERVATOIRE DE PARIS - STATION DE RADIO-ASTRONOMIE DE NANCY, NANCY

²LABORATOIRE PRISME, UNIVERSITÉ D'ORLÉANS, SITE GALILÉE, ORLÉANS

Introduction

La sensibilité des radiotélescopes actuels est de dix ordres de grandeur plus importante que celles des systèmes de télécommunications modernes. En effet, les temps de pose sur le ciel peuvent être très longs (plusieurs heures à comparer avec les fractions de seconde pour les systèmes de télécommunication) et les surfaces collectrices très importantes (le futur radiotélescope international SKA, Square Kilometre Array, de par sa surface sera 10 à 100 fois plus sensible que les radiotélescopes actuels). De plus, certains récepteurs disposent d'une cryogénie pour limiter le bruit système. Tous ces aspects techniques, justifiés par des besoins scientifiques, contribuent à rendre les radiotélescopes de plus en plus performants mais également de plus en plus susceptibles aux interférences radioélectriques (RFI, Radio Frequency Interferences). Or, dans le même temps, la pression des utilisateurs sur le spectre électromagnétique rend l'occurrence de RFI dans les observations astronomiques de plus en plus probable.

Dans le cadre du projet européen Radionet FP7, la station de radioastronomie de Nancy, en collaboration avec l'université d'Orléans, travaille sur l'implantation

d'algorithmes de traitement en temps réel des interférences sur une carte de calcul numérique universelle, Uniboard (UNiversal BOARD for radio astronomy). Dans cet article, nous commençons par présenter deux techniques de détection de RFI qui y seront implantées : détecteur d'impulsions radar et détecteur cyclostationnaire. Pour finir, nous décrivons l'implantation de ces algorithmes dans ladite carte électronique. Ces résultats seront utilisés pour la détection des signaux émis par les pulsars qui, outre la dispersion liée à leur propagation, sont altérés par les RFI (voir encadré).

Détecteurs d'interférences radioélectriques

De nombreuses techniques de traitement des RFI ont été proposées [1]. Pour cette application, nous nous sommes focalisés sur deux types de RFI : les RFI « large bande » et les RFI « bande étroite ». La première technique permet de supprimer, dans le domaine temporel, des interférences très courtes générées par des systèmes radar. La deuxième technique exploite une périodicité cachée présente dans les signaux de télécommunication reçus. Les canaux fréquentiels exploitables pour la radioastronomie sont ainsi sélectionnés automatiquement. Cette dernière approche est partiellement similaire aux techniques développées en télécommunication pour la radio cognitive.

- Détecteur de RFI impulsif de type radar

Pour la détection en temps réel des impulsions radar sans information *a priori*, un détecteur robuste

ABSTRACT

Radio astronomical pulsar observations require specific instrumentation and dedicated signal processing to cope with the dispersion caused by the interstellar medium. Moreover, the quality of observations can be limited by radio frequency interferences (RFI) generated by Telecommunications activity. In this paper, we describe two real time algorithms proposed to remove RFI. First method, designed for rather strong RFI, is based on robust power threshold detection. The second one is based on the RFI cyclostationary properties. These algorithms have been implemented in a high performance dedicated backend UNIBOARD.

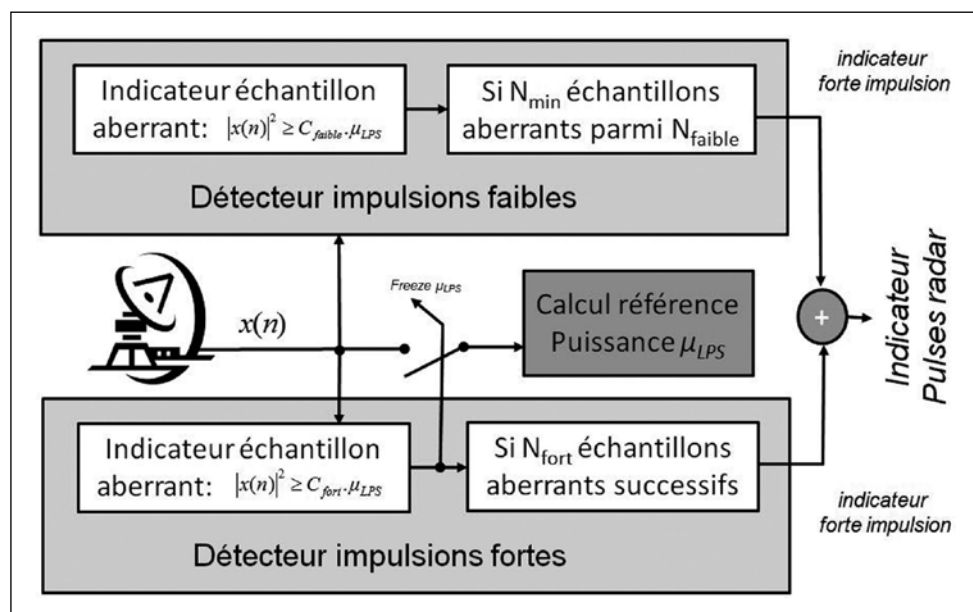


Figure 1: Schéma de principe du détecteur d'impulsion radar.

a été proposé dans [2]. De manière simpliste, la tâche du détecteur d'impulsions radar consiste à détecter, dans un bloc de N échantillons, ceux qui dépassent un certain seuil. Cependant, l'efficacité de la détection dépend du choix du seuil de décision. Ce dernier est déterminé par un algorithme robuste qui limite les fausses alarmes. La figure 1 présente les différents blocs du détecteur : estimation robuste de la puissance de référence, détection d'impulsions fortes, détection d'impulsions faibles.

En radioastronomie, la statistique des signaux naturels observés suit une loi gaussienne centrée. La statistique de leur puissance instantanée suit donc une loi du χ^2 à deux degrés de liberté. Seule sa moyenne est nécessaire pour définir la statistique du signal. Cette valeur moyenne de la puissance de référence, notée μ_{PLS} , est estimée de façon récursive grâce à un filtre de type passe-bas récursif

$$\mu_{LPS}[n] = \mu_{LPS}[n-1] + \alpha (|x(n)|^2 - \mu_{LPS}[n-1]) \quad \text{avec } \alpha < 1 \quad (1)$$

La constante α est équivalente à un facteur d'oubli. Nous avons choisi α de façon à assurer une constante de temps d'établissement du système légèrement inférieure à la milliseconde. À titre de comparaison, la période typique des radars rencontrés est de 0,67 ms. De plus, pour rendre cet estimateur robuste, les échantillons aberrants ne sont pas utilisés pour la mise à jour. Ces derniers sont détectés grâce au détecteur ci-après.

Pour des impulsions radar fortes, un premier seuil est fixé à $S_{fort} = C_{fort} \mu_{PLS}$ avec C_{fort} une constante définie par l'utilisa-

teur (habituellement $C_{fort} = 4$). Tout dépassement ponctuel de ce seuil déclenche, via le signal, *Freeze μ_{LPS}* , le gel de la mise à jour de la puissance de référence, μ_{PLS} . Quant à la détection d'un pulse radar, elle est positive si N_{fort} échantillons consécutifs dépassent le seuil fixé.

Un second étage de détection permet de s'attaquer aux impulsions radar faibles. Comme précédemment, le seuil est calculé à partir de la puissance de référence, μ_{PLS} : $S_{faible} = C_{faible} \mu_{PLS}$. Par rapport au détecteur précédent, la différence réside dans les choix, respectivement, de la constante C_{faible} , d'une fenêtre temporelle plus large de N_{faible} échantillons et d'une condition de déclenchement spécifique. Ainsi, la présence d'une impulsion faible est validée lorsqu'un nombre minimal de détections, N_{min} , a eu lieu dans cette fenêtre temporelle.

Les deux détecteurs d'impulsions fortes et faibles sont ensuite fusionnés pour fournir une décision finale sur la qualité des N_{fort} ou N_{faible} échantillons. À la figure 2 (a), nous présentons une partie des observations menées sur le radiotélescope décimétrique de Nançay (NRT) avec et sans élimination des impulsions radar. Il apparaît qu'une fois détectées et éliminées, les interférences radar n'altèrent plus la ligne de base de l'observation.

- Détecteur cyclostationnaire

La plupart des signaux de télécommunication présentent une périodicité cachée en raison des caractéristiques périodiques impliquées dans la construction du signal (fréquence porteuse, vitesse de transmission, ...). Ces paramètres sont généralement brouillés et cachés par le caractère aléatoire du message à transmettre. Dans le cadre de notre travail nous

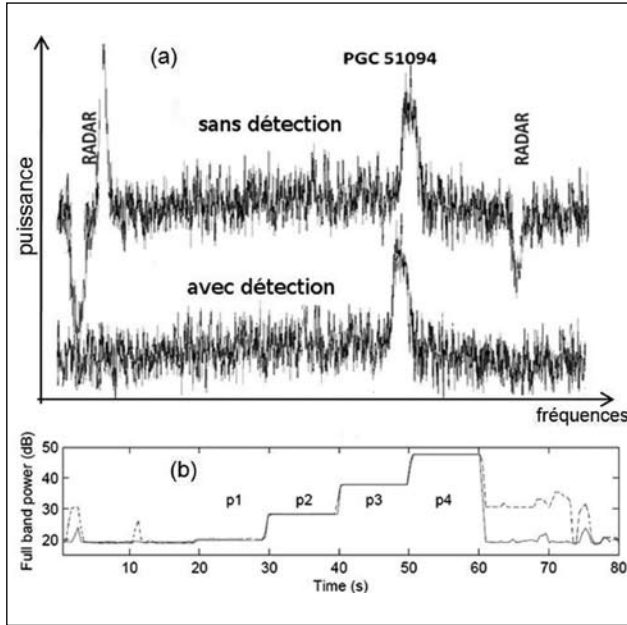


Figure 2 : (a) détecteur de radar, exemple de la raie HI de PGC 51094 observée au NRT avec et sans élimination radar. La bande observée est de 14 MHz et les paramètres du détecteur sont $N_{\text{fort}} = 3$, $C_{\text{fort}} = 4$, $N_{\text{faible}} = 30$, $N_{\text{min}} = 25$, $C_{\text{faible}} = 0,8125$ (b) détections de RFI cyclostationnaires obtenues en temps réel dans la bande décimétrique. Quatre pulses (p1 à p4) issus d'une diode de calibration sont présents. Ils simulent une source astrophysique. La courbe pointillée est la puissance mesurée sans traitement (i.e. avec RFI), la courbe pleine est la courbe après élimination des RFI. Seuls les RFI sont détectés. Les pulses simulant la source astrophysique, malgré leur puissance, ne sont pas éliminés.

nous sommes intéressés au détecteur cyclostationnaire basé sur le critère défini par :

$$D_N^\alpha = \frac{\sqrt{N} C_N^\alpha}{C_n^0} \quad \text{avec} \quad C_n^\alpha = \frac{1}{N} \sum_{n=0}^{N-1} |s(n)|^2 e^{-j2\pi\alpha n} \quad (2)$$

où α est la fréquence cyclique représentant cette périodicité cachée et $s(n)$ est le signal observé sur N échantillons. Ce détecteur repose sur une recherche de périodicités dans les fluctuations de la puissance instantanée. Le détail des performances de ce détecteur est donné dans [3]. La fréquence cyclique α n'étant généralement pas connue, une transformée de Fourier permet de tester rapidement N valeurs de régulièrement espacées.

La figure 2 (b) montre un résultat obtenu en temps réel dans la bande décimétrique à l'observatoire de Nançay. Les RFI sont réels et le signal astrophysique est simulé par des pulses issus d'une diode à bruit. La bande totale est de 7 MHz. Le nombre de canaux fréquentiels est de $N = 2\,048$. Pour chaque canal, $N = 2\,048$ échantillons sont utilisés pour calculer le critère défini à l'équation (2).

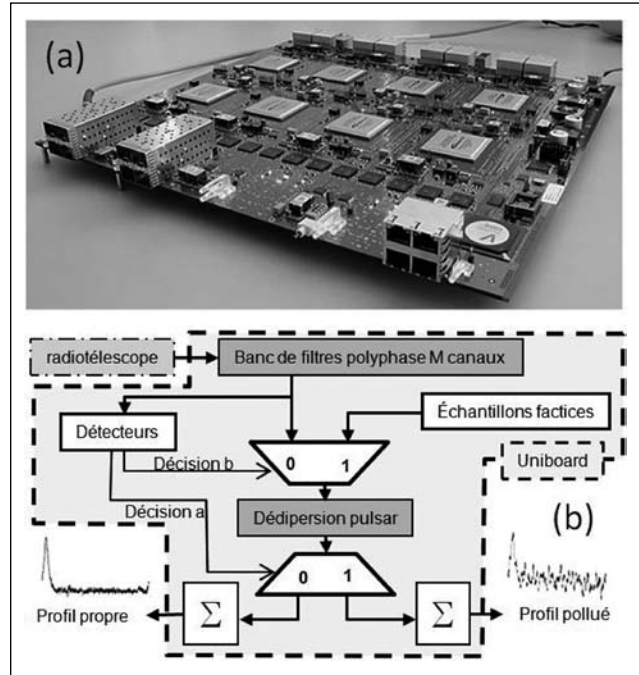


Figure 3 : Le projet UNIBOARD - (a) : La carte UNIBOARD, développée par les instituts partenaires Astron et Jive, est constituée de huit composants programmables (FPGA) de la société Altera (Altera Stratix IV EP4SGX230KF40). Chaque FPGA possède quatre liaisons 10 GbitE pour se connecter directement vers l'extérieur grâce à des interfaces à liaison optique ou cuivre, soit un débit total 160 Gbit/s pour chaque colonne. Chaque FPGA dispose également de deux banques de mémoires DDR3 de 4 Go - (b) Présentation des différentes fonctions implantées dans la carte dans le cadre de l'application « récepteur de signaux pulsar ». Cette application est réalisée en partenariat avec l'université de Manchester.

Implantation sur la carte UNIBOARD

La figure 3 présente la carte UNIBOARD et donne l'ensemble des fonctions implantées dans le cadre de l'application « récepteur de signaux pulsar », du banc de filtres polyphasé à l'intégration des profils pulsar, en passant par le dédisperseur (voir l'encart sur les pulsars). Les différents blocs liés à la détection et à la gestion des RFI sont également insérés.

Notamment, une fois que les RFI ont été détectées, il faut prendre l'action appropriée pour garantir la qualité des observations obtenues. Deux options sont proposées :

- Soit fournir un signal (décision a) qui indique au récepteur la présence d'échantillons pollués. Ces échantillons sont alors mis de côté pour une analyse ultérieure. Les échantillons peuvent même être classés par niveau de puissance. Cette option évite toute perte d'information pour le radioastronome. Il pourra décider en temps différé d'utiliser ou non les données supposées polluées.
- La seconde possibilité est plus radicale (décision b). Il s'agit de remplacer en temps réel les échantillons corrompus par des échantillons factices.

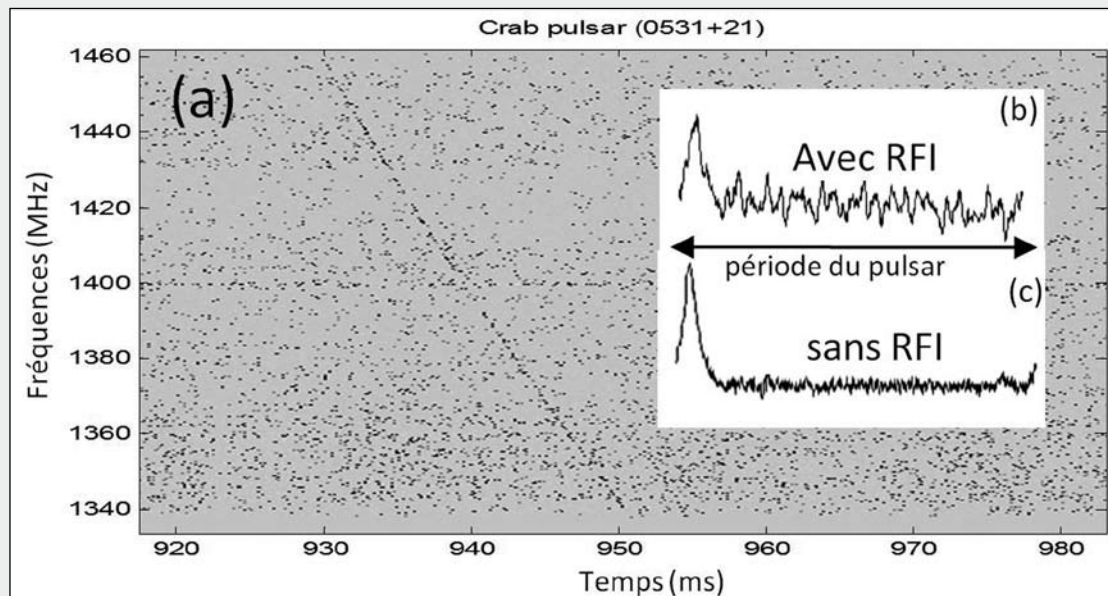


Figure 4 : Représentations temps et fréquence d'un pulse d'un pulsar situé dans la constellation du Crabe (a) Représentation temps fréquence qui illustre la dispersion induite par le milieu interstellaire. (b-c) Profil du pulse obtenu après annulation de la dispersion. La présence de RFI dégrade le rapport signal sur bruit du pulse (b) par rapport au cas (c).

Les émissions des pulsars

Un pulsar est le nom donné à une étoile à neutrons, tournant très rapidement sur elle-même (période typique de l'ordre de la seconde à quelques millisecondes) et, émettant un fort rayonnement électromagnétique qui balaie au cours du temps un cône du fait de la rotation de l'astre. Pour un observateur distant, un pulsar se signale sous la forme d'un pulse périodique, la période correspondant à la période de rotation de l'astre. Cependant, la propagation à travers le milieu fortement ionisé crée, outre une très forte atténuation, un étalement temporel du pulse, c'est le phénomène de dispersion (cf. Figure 4). La détection de ces pulses nécessite donc de mettre en œuvre une instrumentation spécifique qui va permettre d'inverser le processus de dispersion [4]. La présence d'interférences radioélectriques est évidemment une source potentielle de dégradation de la qualité des observations obtenues.

L'option « a » peut être cumulée avec l'option « b ». Cela permet de conserver une trace des modifications imposées au signal.

La nature des échantillons factices peut être définie suivant plusieurs possibilités :

- Échantillons simplement remplacés par des zéros (couper). Son usage doit donc être limité aux cas où seulement quelques échantillons sont pollués, comme par exemple dans le cas du radar.
- Échantillons choisis aléatoirement dans une zone propre du signal observé (copier + coller aléatoire). Ce dernier point est réalisé grâce à une mémoire circulaire contrôlée par le détecteur. Elle contient les N derniers échantillons non pollués de chaque canal.

Conclusion

Deux techniques de détection de RFI ont été présentées dans cet article. L'intérêt de ces techniques a été prouvé dans

[4] afin d'observer des pulsars. Les résultats obtenus nous ont motivés à étendre ces techniques à d'autres domaines d'applications astrophysiques puis de les implémenter sur la carte UNIBOARD. En outre, la station de radioastronomie et l'université d'Orléans travaillent également sur d'autres algorithmes exploitant les possibilités de diversité spatiale [5] qu'offrent les radiotélescopes de nouvelle génération tels que SKA ou LOFAR.

D'un point de vue général, il est important de rappeler que l'efficacité de tout traitement des RFI est limitée et que la sensibilité d'un radiotélescope est toujours dégradée lorsque de telles techniques doivent être mises en œuvre. Il ne faut donc pas interpréter ces possibilités de traitement comme des ouvertures vers un assouplissement des normes d'émission dans les bandes radio. A ce titre, une des pistes à privilégier en priorité est bien évidemment la mise en place de zones de silence radio autour des radiotélescopes. C'est le

moyen le plus efficace de garantir le maximum de sensibilité et de limiter les coûts financiers gigantesques qu'induisent forcément le déploiement de calculateurs numériques avec leurs algorithmes d'élimination.

A terme, il est envisagé d'utiliser une batterie de détecteurs qui alimenteront une base de données sur la qualité des observations. Le radioastronome pourra alors consulter cette base pour vérifier l'intégrité de ses données avant leur exploitation scientifique

Remerciements

Les auteurs remercient la Commission européenne du FP7 [projet PrepSKA (contrat n° 212243) et le projet Radionet FP7 UNIBOARD (contrat n° 227290)], l'ANR (contrat n° ANR-09-BLAN-0225-04) pour le financement d'une partie de ce travail.

Références

- [1] A.-J. Boonstra, R. Weber, "RFI Mitigation Methods Inventory", SKADS, FP6 European project, DS4T3 Deliverable 1, 72 p., 2009.
- [2] C. Dumez-Viou, "Radioastronomical sources restoration from hostile radioelectric environment: Implementation of real-time detectors for dynamic spectra analysis" Ph. D thesis, université d'Orléans, September 2007.
- [3] R. Weber, P. Zarka, V.-B. Ryabov, R. Feliachi, J.-M. Grießmeier, L. Denis, R.-V. Kozhyn, V.-V. Vinogradov, P. Ravier, "Data preprocessing for decametre wavelength exoplanet detection: an example of cyclostationary rfi detector", EUSIPCO'07, Poznan, Poland, September 2007.
- [4] D. Ait-Allal, R. Weber, I. Cognard, G. Desvignes, G. Theureau, "RFI mitigation in the context of pulsar coherent de-dispersion," EUSIPCO, Glasgow, August 2009.
- [5] G. Hellbourg, R. Weber, C. Capdessus, A.-J. Boonstra "Cyclostationary approaches for spatial RFI mitigation in radio astronomy", Comptes-Rendus Physique, 13, 2012.

LES AUTEURS

RODOLPHE WEBER, diplômé de l'Ecole Nationale Supérieure de l'Electronique et de ses Applications à Cergy Pontoise, est Maître de Conférences à l'Ecole Polytechnique d'Orléans. Spécialisé en traitement numérique du signal, il est très impliqué dans des projets de radio télescopes de nouvelle génération (projet LOFAR et SKA) à la fois sur des aspects « architectures de cartes numériques » et sur des aspects algorithmiques pour l'élimination des interférences radioélectriques.

CÉDRIC DUMEZ-VIOU est ingénieur de recherche. Ses travaux ont débuté en 2003 en tant que doctorant sur l'étude et l'implémentation d'algorithmes temps réel implémentés en FPGA, DSP et PC. La finalité de telles études portait, d'une part, sur le traitement des interférences rencontrées lors d'observations radioastronomiques en bande décimétrique (Soleil et Jupiter) et en bande décimétrique (raies HI et OH), et d'autre part, sur la détection automatique par critère morphologique des émissions du couple Io-Jupiter en milieu radio perturbé. Depuis, 2009, Cédric Viou est ingénieur de recherche CNRS à la Station de Radioastronomie de Nançay où il poursuit des développements de récepteurs numériques dédiés à la radioastronomie.

RFI MITIGATION IMPLEMENTATION FOR PULSAR RADIOASTRONOMY

D. Ait-Allal¹, R. Weber^{1,2}, C. Dumez-Viou¹, I. Cognard³, and G. Theureau^{1,3}

¹ Observatoire de Paris

Station de radioastronomie, F-18330 Nançay, France
email: {dalal.ait-allal, Cedric.Dumez-Viou}@obs-nancay.fr

² Institut PRISME, Université d'Orléans

Site Galilée, 12 rue de Blois, 45067 Orléans cedex 2, France
email: Rodolphe.Weber@univ-orleans.fr

³ Laboratoire de Physique et Chimie de l'Environnement et de l'Espace, UMR 6115 CNRS

F-45071 Orléans Cedex 02, France
email: {icognard, theureau}@cnrs-orleans.fr

ABSTRACT

The observation of known pulsars (pulsar timing) or the search for new pulsars can be limited by radio frequency interference (RFI) generated by Telecommunications activity. In this paper we propose several RFI mitigation techniques to cope with impulsive and/or narrow band RFI. For pulsar timing, we have implemented, a pulse blanker and a cyclostationary blanker, both in real time. Example with real data are shown. For pulsar search, we propose a new approach which combines a hardware-efficient search method and some RFI mitigation capabilities. This method is based on a 2D FFT and Radon transform.

1. INTRODUCTION

Pulsars are rapidly rotating highly magnetized neutron stars which produce radio beams that sweep the sky like a lighthouse. The corresponding periodic pulse profiles can be measured with radio telescopes and dedicated backends. Figure 1 describes the typical steps of such a backend. If the pulsar is already known, radio astronomers are interested by its precise timing. A coherent de-dispersion processing is then implemented to compensate the frequency-dependent time-shift due to non-linearities in the interstellar medium (see [1, 7] for more details). If not, an iterative search procedure is implemented to identify new pulsars.

In both cases, the quality of pulsar observations is limited by radio frequency interference (RFI) generated by various (and growing) Telecommunications activities. In this paper, we propose several RFI blanking techniques to cope with impulsive and/or narrow band RFI. Depending on where the detectors are implemented in the system, different types of RFI can be detected.

For pulsar timing, three possibilities have been identified on the system architecture defined on Figure 1:

- a) At the input of the polyphase filter bank, just after digitization. This configuration is appropriate for impulsive or burst broad-band RFI.
- b) Just after the polyphase filter bank.
- c) Just after the FFT. This configuration is more appropriate for narrow band and continuous RFI.

In sections 2 and 3, the implementation of two detectors is described. One is a Pulsed-RFI detection based on power criteria (suitable for pulsar timing case a), the other is a cyclostationary detector dedicated to continuous RFI (suitable for pulsar timing cases b and/or c).

For pulsar search, we propose a new approach which combines a hardware-efficient search method and some RFI mitigation capabilities both for impulsive and narrow band RFI. This method is based on a 2-dimensional FFT (2D FFT) and Radon transform. The algorithm is described in Section 4.

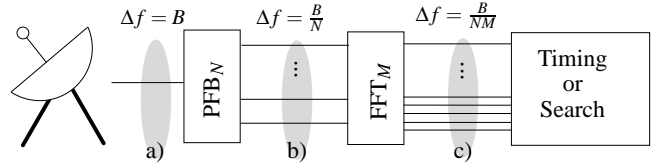


Figure 1: Typical pulsar backend architecture: a) step 1: waveform acquisition b) step 2: polyphase filter bank (PFB) to split the input frequency band (usually hundreds of MHz) into smaller bands (few MHz) c) step 3: During their travel through the interstellar medium, the pulsar pulses are progressively attenuated and spread over time (dispersion effect). Time-frequency planes are calculated for real time de-dispersion procedure in the case of pulsar timing or for specific algorithm in the case of pulsar search (usually an off-line iterative search)

2. PULSED-RFI DETECTION WITH A POWER DETECTOR

The power detector principle is quite simple: outlier samples that exceed a certain level are considered as interference and trigger the blanking of the corresponding data block. However, the efficiency of such detector depends on the robustness of its decision threshold. The issue is that the same block of data is used for threshold estimation and for outlier detection. Moreover, in terms of logic resources and data bus width, the implementation cost of such an approach could be an issue (see Hampson [6]).

2.1 Threshold calculation

Our approach is based on the χ^2 distribution model. So, only one parameter (the mean μ) is needed to fully define the signal statistics. Consequently, the threshold value S is calculated as $S = C\mu$ with C a parameter defined by the user. The mean value is recursively estimated (see Fig. 2-step 1). To make this estimation robust against outliers due to RFI, the strongest samples are systematically discarded. The effect of this clipping on mean estimation can be theoretically derived (Dumez-Viou [5]) and the proper correction is included in the parameter C . Compared to classic power detectors, two improvements are proposed to enhance the detection performances for strong and weak radar pulses. They are based on the temporal characteristics of the expected RFI, here radar pulses.

2.2 False alarm rate reduction

For synchronisation reasons, our system blanks the whole block of data even if only one outlier has been detected in that block. This

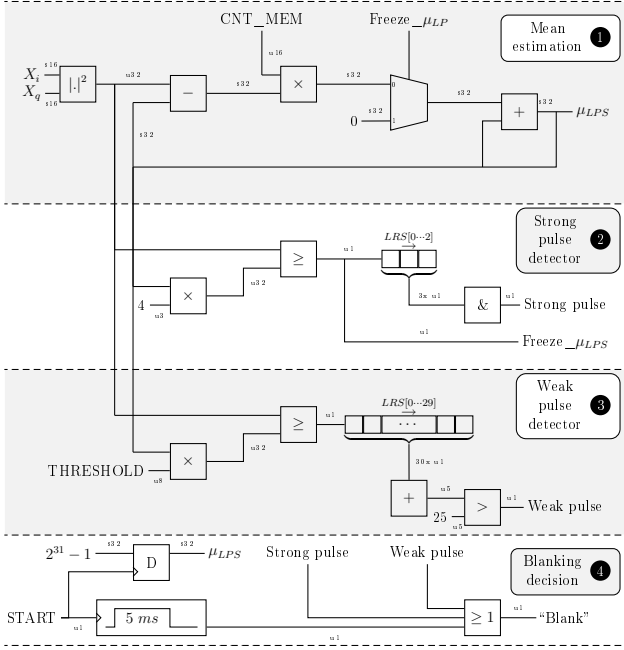


Figure 2: Hardware for the radar pulse blander. LRS is a Logical Right Shift unit. $s..$ and $u..$ specify the databus widths for signed and unsigned operands. The implementation is composed of 4 blocks: 1) recursive robust mean estimation 2) Strong pulse detector 3) Weak pulse detector 4) decision level for blanking.

approach increases the false alarm rate. For example, if the threshold is set to 4μ to produce a 2.75% false alarm rate, a 2048-sample block configuration will systematically trigger the detector and the receiver returns no usable results. To overcome this problem, note that the time distribution of such false alarms over the whole observation is uniform. Let α be the false alarm rate. Then, the probability of obtaining N consecutive uncorrupted samples that trigger the detector is α^N .

For strong radar pulses, let us consider the following parameters: $S = 4\mu$, $N = 3$, $\alpha = 2.75\%$. Thus, triggering the blander if three consecutive samples exceed the threshold sets a new false alarm rate of $(2.75\%)^3 = 2.1 \times 10^{-3}\%$ leading to the blanking of only 6.4% of the data blocks. Besides, the hardware required to handle the previous modifications is limited to a one-bit wide $(N - 1)$ -bit deep shift register to store the results of the comparisons between samples and threshold, and a N -bits AND gate (see Fig. 2-step 2).

For weak radar pulses, considering N consecutive samples leads to poor detection performances, since radar bursts are increasingly buried under the system noise. However, tests have shown that counting the number of detections in a time window is better. Thus, for our weak pulse detector, we set the threshold to 0.8125μ , the time window to 30 samples and the triggering number of detection to 25. Using a binomial law and a tabulated χ^2 distribution, a false alarm rate of $1.2 \times 10^{-3}\%$ is achieved, resulting in the blanking of 4% of the data blocks. In terms of implementation, rather than adding the 30 bits for each new sample, we recursively calculate the sum. The balance (-1, 0 or +1) of '1' entering and '1' leaving the shift register is added to the previous count. The hardware now extends to a one-bit wide 30-bit deep shift register, a 1-bit subtractor and a 2-bit/5-bit adder (see Fig. 2-step 3).

2.3 Conclusion

To improve the detection performance, we focused on the pulse length, which is one of the basic characteristics shared by many radar systems. The choice of $N = 3$ for strong pulses and $N = 30$ for weak pulses has been done experimentally. The performances

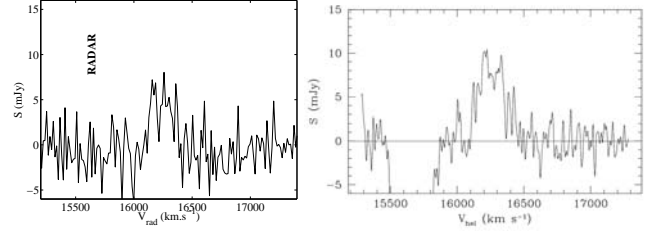


Figure 3: HI line of ESO 474-G26. Right: Observed at the Nançay Radio Telescope (NRT, [11]) with pulse blanking. left : Observed at NRT without pulse blanking.

of our detectors are shown in Fig. 4. The design operates at a maximum sampling rate of 145 Ms/s. The logic gates used to implement the algorithm occupy 4% of a 3 Mges FPGA (Virtex II from Xilinx) and 2% of the 96 18×18 -multipliers available. The design has been used to observe cosmic sources with flux densities as low as 5 mJy. No radar residuals could be seen on the base line as shown on figure 3.

3. CYCLODET: A CYCLIC DETECTOR

It is also possible to apply power detection on narrow band RFI. However, precise thresholding implies careful calibration of the frequency band and it assumes that the background power level of the signal is constant over time.

A better approach is to use a criterion that can differentiate quasi-steady noise-like signals from artificial or quickly varying signals. Cyclostationarity is such a criterion,[10]. Most telecommunication signals present a hidden periodicity due to the periodic characteristics involved in the signal construction (carrier frequency, baud rate, coding scheme...). These parameters are usually scrambled and hidden by the randomness of the message to be transmitted. In [1], we have demonstrated the interest of such an approach for pulsar applications. In the following, the detector principle is briefly presented and its real time implementation is described.

3.1 Principle of cyclostationary detection

We assume that $x(t)$ is a mix of a stationary signal (*i.e.* a cosmic source and/or the system noise) and a cyclostationary signal (*i.e.* a RFI). Let us consider the following criterion:

$$C_N^\alpha = \frac{1}{N} \sum_{n=0}^{N-1} |x|^2(n) \exp(-j2\pi\alpha n) \quad (1)$$

where α is the cyclic frequency. This parameter is linked to the above-mentioned periodic characteristics. This detector expresses the search for periodicities in the instantaneous power fluctuations. To make this detector robust against slow power variations, we define a normalized version of our previous criterion:

$$D_N^\alpha = \frac{\sqrt{N}C_N^\alpha}{C_N^0} \quad (2)$$

In [12], we have derived the statistical properties of this detector as a function of the interference to noise ratio (INR).

3.2 Hardware implementation

From the above consideration, an operational cyclic detector has been implemented on a real time digital backend at Nançay Observatory. The algorithm is implemented into a digital programmable component Virtex II, an FPGA from the Xilinx company. The successive steps are (see also figure 5):

1. Channelization of the signal coming from the radio telescope. The signal in each channel is assumed to be complex. This process is done in real time by the digital receiver (*i.e.* output b or c in figure 1).

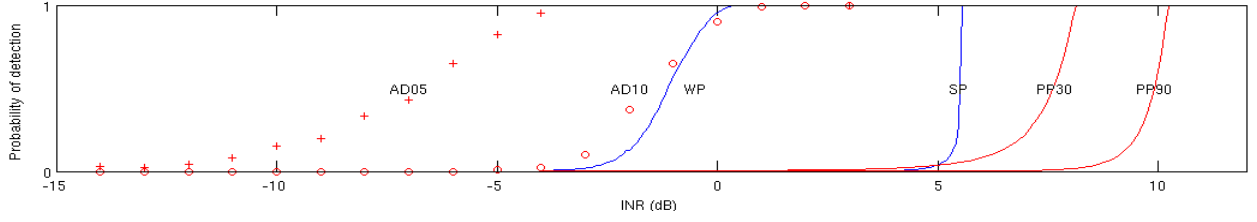


Figure 4: Probability of detection of several detectors as a function of radar pulse Interference to Noise Ratio (INR). The blue curves annotated SP (respectively, WP) correspond to the behavior of the strong pulse detector (respectively, weak pulse detector). PP30 and PP90 refer to [9]. They correspond to simple pulse detectors for two different settings. AD05 and AD10 refer to [4]. They are based on a much more advanced algorithm that gives better performances for its finest setting. However, it includes a lot of information about a specific radar pulse shape. Any radar pulse whose shape differs significantly from the model will not be detected as easily.

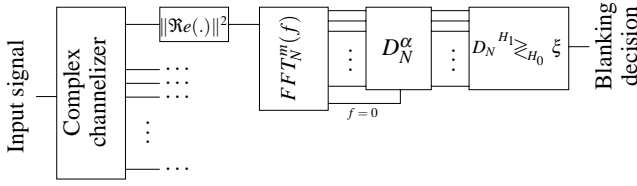


Figure 5: Cyclostationary detector algorithm

2. To reduce the computational load of the cyclic detector, the algorithm is applied to the real part only ($s_r(n)$) of the signal. We compute the Fourier transform, $FFT_N^m(f)$, over N samples on $s_r^2(n)$ for the m^{th} channel, $m = 1, \dots, M$. M is the number of channels.
3. According to a given threshold ξ derived from the theoretical study ([12]), we will consider that an RFI is present on the m^{th} channel if:

$$\exists k > 0 / \frac{\sqrt{N} |FFT_N^m(k)|}{FFT_N^m(0)} \geq \xi \quad (3)$$

with $\xi = \sqrt{-2 \log(p_{fa})}$

where p_{fa} is the expected probability of false alarm.

The design can process a 7 MHz frequency band. The logic gates used to implement the complex channelizer and the cyclic detector occupy 80% of the slices of a 3 M gates FPGA (Virtex II from Xilinx), 87% of the embedded memory and 88 of the 96 18×18 multipliers available. An external memory based on a SDRAM is also required to store a buffer of 2048x2048 16-bits complex-words (16 Mbytes).

Figure 6 shows some results obtained in the decameter band where calibration signals (wide band stationary noise) were used to simulate four pulsar pulses with different power levels. The red time-frequency slots correspond to RFI detected by our real time detector. Narrow band RFI are clearly detected. The detector is also sensitive to wide band RFI, and to noise diode power level transitions. The impulse or the step of power generated by these events produces a high frequency content in the cyclic spectra that triggers the cyclostationary detector. However, when the noise power returns to a steady state, the detector is again blind to power levels. For pulsar observations, these experiments demonstrated that the cyclostationary detector can discriminate between bursts due to pulsars and bursts due to RFI.

4. REAL TIME PULSAR SEARCH PROCEDURE

The search for new pulsars is a difficult task since neither the dispersion measure (DM) nor the periodicity are known [7]. In figure 7.a, the classical search procedure is described. It consists in recording

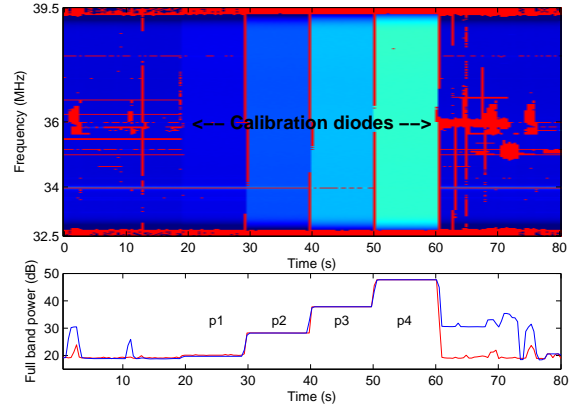


Figure 6: Results obtained in the decameter band at Nançay 1/2 observatory. (top) Time-frequency power plane. The red time-frequency slots correspond to RFI detected by our real time detector. (bottom) Total power versus time. The red curve is the total power after blanking. The total bandwidth is 7 MHz. The number of channels is $M = 2048$. On each channel, $N = 2048$ samples are used to compute the criterion. The cyclic detector is insensitive to steady power levels generated by the successive calibration noise diode pulses (p1 to p4) added to the signal.

all the data and in incoherently de-dispersing them by trying off-line a set of DM. Then, it can be searched for periodicity by using the fourier transform. Finally, all the data are folded with different set of periods, providing a list of pulsar candidates. This approach needs huge disk space and is quite time consuming. In this section, we propose a new real time pulsar search method, requiring less disk space, fewer computational resources and providing a simple way to blank RFI. It is based on a two dimensional Fourier transform (2D-FFT) and Radon transform. 2D-FFT is mentioned in the pulsar literature but only as a tool to search for periodicity ([2], [8]).

4.1 2D-FFT method

In figure 7.b, the different steps of the algorithm are presented and in figure 8 an example with a real pulsar pulse is provided for illustration. In short :

- The final output of the real time part consists of the $N \times N$ image corresponding to the mean of M consecutive $|2D - FFT|^2$ images. An example of such an image is given at figure 8.b. Each 2D-FFT is computed on independent $N \times N$ time-frequency power planes such as the one provided at figure 8.a. With this approach, all dispersed pulsar pulses are projected at the same location in the 2D-FFT image. This location depends only on

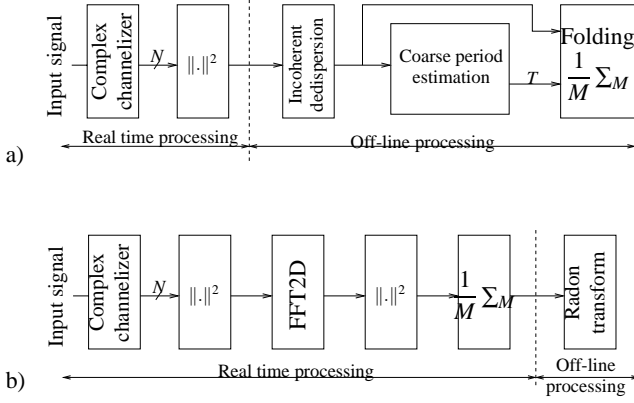


Figure 7: Different pulsar search procedures: (a) the classical approach. The output of the N -channel filter bank is stored. Off-line, different pulsar dispersion measures are used to de-disperse the pulsar pulses. Then, different periodicities are iteratively used to fold the successive pulsar pulses. (b) The 2D-FFT approach. A 2-dimensional FFT is applied on the output power of the N -channel filter bank providing an $N \times N$ image. The average of M images are stored. Then, a Radon transform is applied.

the pulsar DM and is different from the one obtained for RFI. In other words, any impulsive (respectively narrow band) RFI will be concentrated in the horizontal (respect. vertical) line centered in the 2D-FFT image. Thus, to remove RFI and to obtain a clean image, all that is required is to blank these vertical and horizontal lines which cross at the center of the 2D-FFT image.

- Off-line, a Radon transform is applied on the cleaned 2D-FFT image. It consists in summing the image intensity along successive radial lines [3]. When this integration line corresponds to the line where the pulsar pulses have been projected, all the pulsar power is concentrated in one point. At other integration angles, only the noise contribution will be integrated. Figures 8.c and 8.d illustrate this principle with respectively the dirty and clean image.

The advantage of this approach is that only a limited amount of data (the $N \times N$ 2D-FFT image) is stored for off-line processing. Furthermore, it provides simple RFI mitigation capabilities. In the following, we investigate the sensitivity of this approach compared with the classical one.

4.2 sensitivity comparison

We consider the following signal model and algorithm parameters:

- The background noise is a white zero-mean Gaussian noise with power σ_n^2 .
- The pulsar pulse is a white zero-mean Gaussian noise with power σ_p^2 . The pulse duration is L . The signal to noise ratio is defined by $SNR = \sigma_p^2 / \sigma_n^2$. For the theoretical sensitivity derivation, we assumed that $\sigma_p^2 \ll \sigma_n^2$.
- There is no RFI.
- The t-f power plane dimension is $N \times N$.
- For the classical approach, we consider that the dispersion measure (DM) and the pulsar periodicity has been found by the ad-hoc iterative procedure. Thus, it is equivalent to assume 1) DM=0 (i.e. vertical pulse) and 2) pulses always occur at the same position in the t-f power plane. These t-f power planes are integrated over frequency to provide individual pulse profiles. Finally, M pulse profiles are summed together. The sensitivity, $S_{classic}$ is defined by the average profile deviation due to the pulse over the noise only profile standard deviation.
- For the 2D-FFT approach, we set DM to zero as well. Indeed, with no loss of generality, we consider that zero-DM is just a

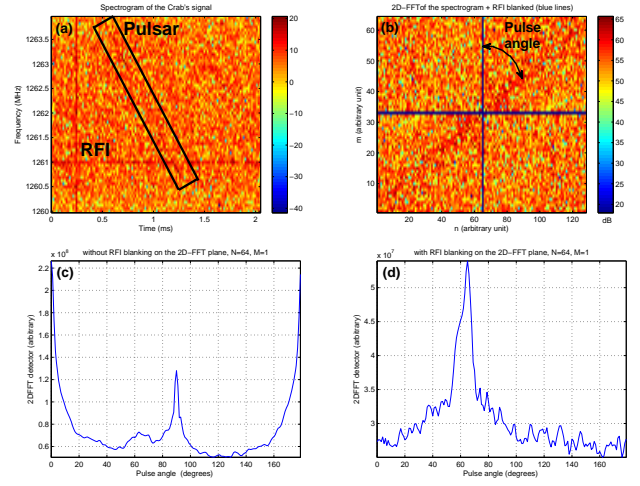


Figure 8: Application of 2D-FFT pulsar search approach on the Crab pulsar. (a) 128×64 time-frequency power plane of the Crab pulse. We added a couple of synthetic RFI. (b) 2D-FFT image time-frequency power plane. The blue cross in the middle corresponds to the blanked RFI locations. (c) Radon transform of the dirty 2D-FFT image (i.e. RFI locations not blanked). The pulsar detection is not possible. (d) Radon transform of the clean 2D-FFT image. The pulsar can be easily detected.

DM among others. The central pixel of the 2D-FFT image is systematically blanked. The average of M 2D-FFT planes provides the final clean 2D-FFT image. Finally, the Radon transform is applied for different angles. The sensitivity, S_{2DFFT} is defined by the Radon transform deviation at angle 90° over the Radon transform standard deviation at other angles.

This model is illustrated in figure 9. In particular, output examples of each approach are proposed and a graphical definition of sensitivity is provided. By considering the mean and the variance of this model at the successive algorithm stages, we have derived the following theoretical equations:

- for the classical case: $S_{classic} = \frac{\sigma_p^2}{\sigma_n^2 / \sqrt{NM}} = SNR \sqrt{NM}$
- for the 2D-FFT case: $S_{2DFFT} = \frac{T \sigma_p^4}{\sigma_n^4 / \sqrt{NM}} = T \cdot SNR^2 \sqrt{NM}$ where $T = L/N$ (i.e. T is pulse width in the t-f power plane).

In figure 10, these sensitivities have been computed by simulation for different parameters. Theoretical sensitivity are also plotted. From this figure, several remarks can be stated:

- It is shown that theory and simulations fit together. The shift for ($SNR = 0$ dB, $L = 2048$) is due to the fact that our assumption ($\sigma_p^2 \ll \sigma_n^2$) is no longer valid.
- The 2D-FFT sensitivity is proportional to SNR^2 . This makes this approach less sensitive to low SNR compared with the classical approach (see for example the case ($SNR = -5$ dB, $L = 256$)). This drawback is counterbalanced by the fact that the 2D-FFT approach is also sensitive to the pulse width, T . The wider the pulse is, the better the 2D-FFT detection is (see case ($SNR = -5$ dB, $L = 2048$)). In some cases, the 2D-FFT approach can outperform the classic method (see case ($SNR = 0$ dB, $L = 2048$)).
- These detection performances are achieved without knowing anything about the pulsar (neither the DM nor the periodicity). This is not the case for the classical pulsar search procedure, where the derived performances are only achieved when exact DM and periodicity are known.

The next step will be to test the algorithm with different configurations of real pulsar data and real RFI.

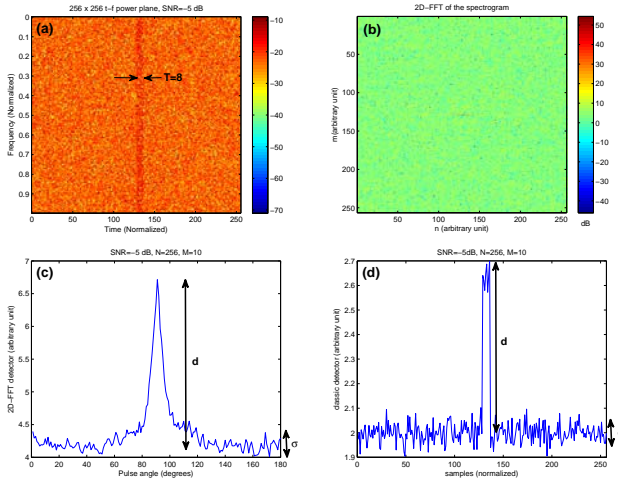


Figure 9: Application of the classical and 2D-FFT pulsar search approach on our zero-DM model. (a) Time-frequency (t-f) plane ($N = 256$) of one pulsar pulse with $\text{SNR} = -5$ dB, $L = 2048$, $N = 256$ and $T = L/N = 8$. (b) Average 2D-FFT image of $M = 10$ t-f planes. (c) Output of the Radon transform. The peak position is directly related to the DM. Here $DM = 0 \Rightarrow \text{angle} = 90^\circ$. (d). Output of the classic approach. For both cases, the sensitivity is defined by the ratio d/σ .

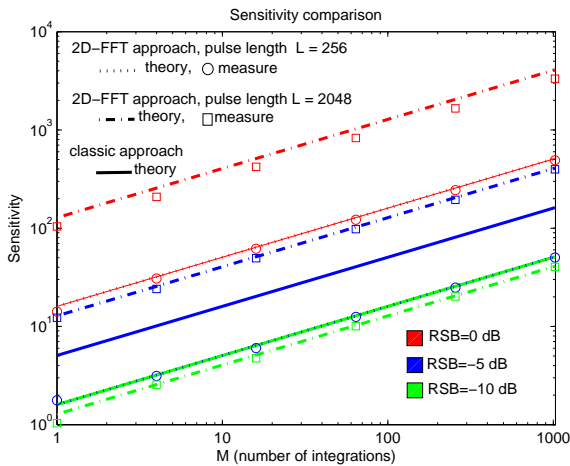


Figure 10: Comparison of the 2D-FFT and the classic pulsar search procedures ($N = 256$). 3 SNR (0 dB, -5 dB and -10 dB) and 2 pulse lengths ($L = 256$ and $L = 2048$) have been simulated for different values of M . The continuous and dashed curves correspond to theoretical values. Box and circles correspond to 2D-FFT sensitivity measurements through simulations.

5. CONCLUSION

Several RFI mitigation techniques applicable to pulsar observation have been presented. In the case of pulsar timing, a Pulsed-RFI detector and a cyclostationary detector have been implemented for real time experiments. In the framework of the UNIBOARD FP7 European project, these algorithms will be implemented in a multi-purpose scalable computing platform for Radio Astronomy as part of the pulsar receiver. In the case of pulsar search, a new approach which combines a hardware-efficient search method and some RFI mitigation capabilities has been proposed. It could be an alternative for the next generation of radio telescopes such as the LOFAR ra-

dio telescope (www.lofar.org) or the Square Kilometer Array (SKA, www.skatelescope.org) where the huge amount of data to be processed by classical search procedures is an issue.

Acknowledgements. The authors would like to thank the European Commission Framework Program 7 (Project PrepSKA (contract no 212243) and Project Radionet/Uniboard (contract no 227290)), The French funding agency ANR (contract ANR-09-BLAN-0225-04) for funding part of this work.

REFERENCES

- [1] D. Ait-Allal, R. Weber, I. Cognard, G. Desvignes and G. Theureau, "Rfi mitigation in the context of pulsar coherent de-dispersion at the Nançay radio astronomical observatory" in *Proc. EUSIPCO 2009*, Glasgow, UK, September 4-8, 2009.
- [2] F. Camilo, D.J. Nice, J.A. Shrauner, and J.H. Taylor, "Princeton-Arecibo Declination-Strip Survey for Millisecond Pulsars. I.," *The Astrophysical Journal*, vol. 469, 1996, pp. 819-827.
- [3] S. Deans, "The radon transform and some of its applications" Krieger Pub. Co., 1983
- [4] W. Dong, B.D. Jeffs and J.R. Fisher, "Kalman Tracking and Bayesian Detection for Radar RFI Blanking," RFI2004, Workshop in Mitigation of Radio Frequency Interference in Radio Astronomy, Penticton, Canada, July 2004
- [5] C. Dumez-Viou, "Radioastronomical sources restoration from hostile radioelectric environment: Implementation of real-time detectors for dynamic spectra analysis," PhD thesis, <http://tel.archives-ouvertes.fr/tel-00319939/fr/>, September 2007
- [6] G.A. Hampson, "Implementation of the Asynchronous Pulse Blanker," June 2002. <http://esl.eng.ohio-state.edu/~rsttheory/iip/docserv.html>
- [7] D. Lorimer and M. Kramer "Handbook of pulsar astronomy," Cambridge (GB): Cambridge University Press, 2005. pp. 106-123, pp. 141-142.
- [8] A. Lyne and F. Graham-Smith "Pulsar astronomy," Cambridge: Cambridge University Press, 2006.
- [9] N. Niamsuwan, J.T. Johnson and S.W. Ellingson, "Examination of a simple pulse-blanking technique for radio frequency interference mitigation," *RADIO SCIENCE*, 40(3), June 2005
- [10] E. Serpedin, F. Panduru, I. Sari, and G.B. Giannakis, "Bibliography on cyclostationarity" *Signal Processing*, vol. 85, pp. 2233-2303, Dec. 2005.
- [11] Station de radioastronomie de Nançay, France, www.obs-nancay.fr
- [12] R. Weber, P. Zarka, V. Ryabov, R. Feliachi, J. Grießmeier, L. Denis, V. Kozhyn, V. Vinogradov, and P. Ravier, "Data preprocessing for decametre wavelength exoplanet detection: an example of cyclostationary rfi detector," *Eusipco*, Poznan, Poland, 2007.

RFI MITIGATION IN THE CONTEXT OF PULSAR COHERENT DE-DISPERSION AT THE NANCAY RADIO ASTRONOMICAL OBSERVATORY

D. Ait allal¹, R. Weber², I. Cognard^{1,3}, G. Desvignes³ and G. Theureau^{1,3,4}

¹Station de radioastronomie de Nançay, Observatoire de Paris, CNRS/INSU Nançay, F-18300 Nançay, France

dalal.ait-allal@obs-nancay.fr

²Institut PRISME – Université d'Orléans 12 Rue de Blois BP 6744 Orléans Cedex 2 F-45067 France,

rodolphe.weber@univ-orleans.fr

³Laboratoire de Physique et Chimie de l'Environnement et de l'Espace, UMR 6115 CNRS, F-45071 Orléans Cedex 02, France,

icognard@cnrs-orleans.fr, Gregory.Desvignes@cnrs-orleans.fr

⁴GEPI, Observatoire de Paris, CNRS, Université Paris Diderot ; Place Jules Janssen 92190

Meudon, France, theureau@cnrs-orleans.fr

ABSTRACT

Radio astronomical pulsar observations require specific instrumentation and dedicated signal processing to cope with the dispersion caused by the interstellar medium. Moreover, the quality of observations can be limited by radio frequency interferences (RFI) generated by Telecommunications activity. This paper will present the innovative pulsar instrumentation based on Graphical Processing Units (GPU) which has been designed at the Nançay Radio Astronomical Observatory. In addition, first simulated results on RFI cyclostationary detectors, which will be implemented on the system, will be described and discussed.

1. CONTEXT

The radio astronomy is a field that uses radio waves emitted by celestial objects for astronomic observations and studies. Among those fascinating objects, pulsars – rapidly rotating highly magnetized neutron stars - produce radio beams that sweep the sky like a lighthouse [1]. If the beam is oriented towards the Earth, it will produce periodic pulses which can be measured with radio telescopes and dedicated backends. During their travel through the interstellar medium (ISM), these pulses are progressively attenuated and spread over time (i.e. pulse higher frequencies will arrive earlier than the pulse lower frequencies). Figure 1.a and 1.b shows a raw time-frequency representation of one of these pulses. This phenomenon, named dispersion, makes these pulses barely detectable without further processing. Moreover, the quality of pulsar observations is also limited by radio frequency interferences (RFI) generated by various (and growing) Telecommunications activities. Figure 2 gives an example of the distortion induced by RFI on pulsar observation.

To invert the ISM effect, specific receiver must be designed, and to limit the impact of RFI, robust signal processing algo-

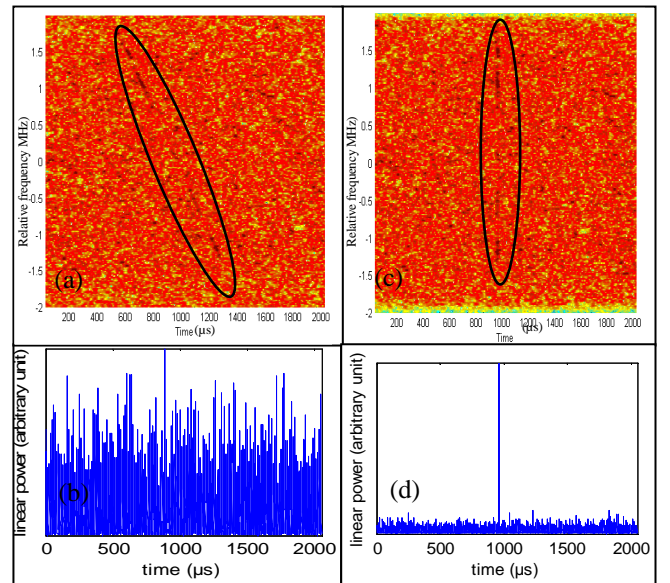


Figure 1–a. Time-frequency representation of a giant pulse from the Crab (PSR 0531+21) as it is measured without further processing. b. Corresponding power fluctuations over time in the 4 MHz band. The pulse is undetectable c. Time-frequency representation of the same giant pulse after de-dispersion. d. Corresponding power fluctuations over time in the 4 MHz band. The pulse is easily detectable.

gorithms must be applied. This paper will present an original backend which has been designed for pulsar de-dispersion and simulations of RFI mitigation techniques which can be applied in such context.

In section 2, the functional description of the coherent de-dispersion algorithm is explained and the innovative hardware architecture of the Nançay de-dispersion backend is detailed. In section 3, a waveform pulsar model is derived. This model will be used to test RFI mitigation algorithms. Their description and their simulation are given in section 4.

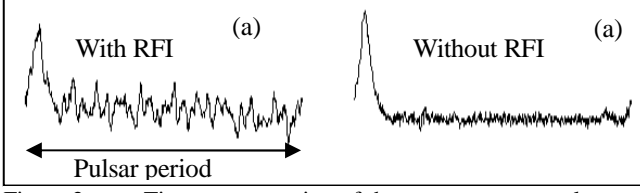


Figure 2 – **a.** Time representation of the average power pulse profile of the pulsar J0034-0721 after coherent de-dispersion but with RFI signals. **b.** Time representation of the average power pulse profile of the pulsar J0034-0721 after coherent de-dispersion and with RFI signals blanked (with other RFI mitigation Scheme).

2. NANÇAY INSTRUMENTATION FOR PULSAR OBSERVATIONS

2.1 Coherent de-dispersion principle

The ISM dispersion can be modelled as a frequency transfer function [1], H :

$$H(f_0 + f) = e^{+j \frac{2\pi DM}{2.41 \times 10^{-10}} \frac{f^2}{(f + f_0)f_0^2}} \quad (1)$$

Where DM is the dispersion measure, f_0 is the center frequency of observed bandwidth Δf and $|f| < \Delta f/2$.

The coherent de-dispersion consists of applying the inverse of this transfer function to the dispersed received signal. This processing is done in the frequency domain by using Fourier transforms. The operations are detailed inside the “Data server” box in Figure 3. The taper function is used to avoid aliasing in the low-pass filtering. The combination of the inverse of the dispersion transfer function and the taper function is known as the chirp function.

2.2 Hardware Implementation

Figure 3 shows the signal dataflow from the radio telescope to the final de-dispersed pulses. The different steps are described below:

- The Analog System [2]:** two orthogonal polarization signals from the radio telescope are downconverted through a custom circuit board quadrature downconverter (QDC). Each QDC handles one polarization and performs a quadrature demodulation providing 128 MHz complex bandwidth which is sent to the SERENDIP5 Spectrometer for digitalisation and channelizing.
- The SERENDIP5 Spectrometer [2]:** SERENDIP5 (Search for Extraterrestrial Radio Emissions from Nearby Developed Intelligent Populations) was designed by CASPER (Center For Astronomy Signal Processing and Electronics Research) for applications in SETI (Search for Extra-Terrestrial Intelligence), Pulsar and general radio astronomy. For this application, it contains four 8-bit analog to digital converters (ADCs) that can be clocked up to 200MHz. A logic programmable device (Virtex 2 XC2V4000 from Xilinx) is used to perform a 32 channel polyphase filter bank (PFB). The channel output bandwidth is 4 MHz complex. An addi-

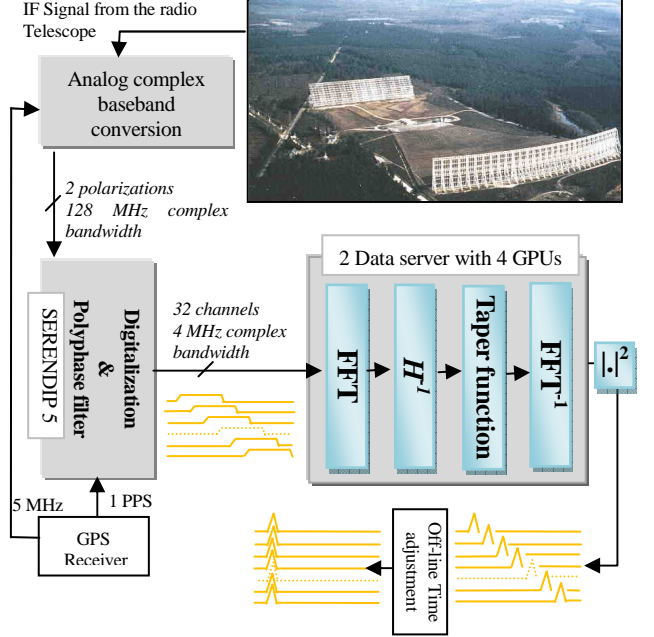


Figure 3 – Description of the instrumentation for pulsar observations which is in operation since July 2008 at Nançay Observatory. The top right hand corner picture is the giant decimetre radio telescope of Nançay.

tional Xilinx device (Virtex 2 XC2V1000) is used as a reconfigurable backend processor which can pass data to an independent computer

- Data servers:** This is the innovative part of the design. The two data servers have two main processing's steps. The first one is to read the raw data in from the SERENDIP5 spectrometer and reorganize them from a channel-order to single-channel time ordered chunks. This step is controlled by four micro processors, that send these chunks to four GPUs [3] (Graphic Processing Unit- NVIDIA 88 GTX ~128 parallel processors) for the final processing's step. The data are then read by the de-dispersion program that performs the following tasks:
 - Convert raw data from binary to floating point.
 - Apply a Fourier transform of appropriate length (8192 bins).
 - Multiply by the de-dispersion filter, H^{-1} and the taper function
 - Inverse Fourier transform back to the time domain.
 - Detect the data to get power versus time, and cross-multiply polarization terms.
- Time adjustment:** offline resynchronisation of all the channels.
- GPS receiver:** It provides SERENDIP5 clock synchronization through a one pulse per second (1PPS) signal and oscillator synchronization through a 5 MHz reference clock.

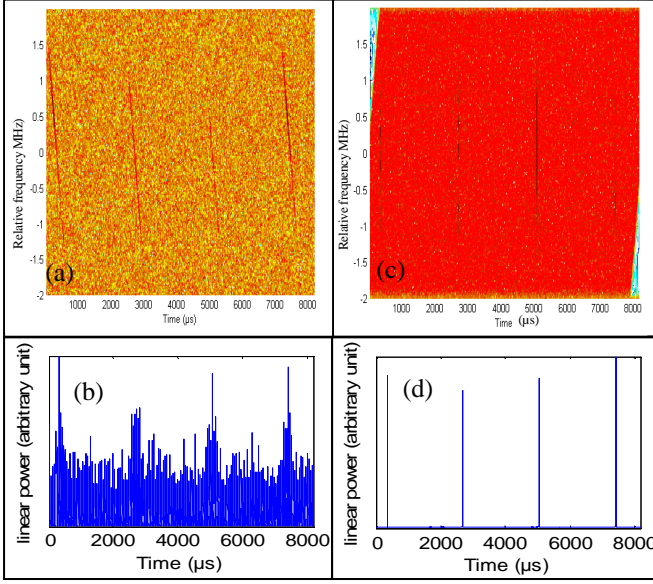


Figure 4 – **a.** Time-frequency representation of the pulsar signal generated by the model given by Equ. 7 ($BW=4\text{MHz}$, $t_{DM}=938.0793\text{ }\mu\text{s}$ and pulsar period is 2.35 ms). **b.** Corresponding temporal power fluctuations over the whole bandwidth. **c.** Time-frequency representation after de-dispersion. Equ. 1 has been used with the following parameters $f_0=1262\text{ MHz}$ and $DM=56.791\text{ pc.cm}^{-3}$. **d.** Corresponding temporal power fluctuations over the whole bandwidth.

This coherent de-dispersion pulsar receiver is operational since July 2008 at Nançay Observatory [4]. Figure 1.c gives the result of this coherent de-dispersion applied on the signal from Figure 1.a. We can see that after the coherent de-dispersion, the pulse is no more spread and its power can be easily detected (Figure 1.d compared to Figure 1.b).

To improve the system sensitivity and its robustness to RFI, it is planned to implement RFI mitigation algorithms in this system. The next sections will describe the preliminary simulated results obtained on that topic.

3. PULSAR SIGNAL MODEL

In our model, we assume that the pulsar beam is represented by periodic Dirac impulsions so that the pulse signal can be modelled as successive chirp functions.

The temporal evolution of the frequency can be written as:

$$f(t) = \frac{\alpha BW}{t_{DM}^2} t^2 - \frac{BW(\alpha+1)}{t_{DM}} t + \frac{BW}{2}, \quad t \in [0, t_{DM}] \quad (2)$$

where BW is the bandwidth (typically 4 MHz), t_{DM} is the dispersion delay over this bandwidth, α is a factor which characterises second order delay distortions. Without loss of generality, all the simulations have been made with $\alpha=0$ (i.e. the frequency becomes a linear function of time). Thus, Equ. (2) becomes:

$$f(t) = -\frac{BW}{t_{DM}} t + \frac{BW}{2}, \quad t \in [0, t_{DM}] \quad (3)$$

Then, the phase of the pulsar signal can be expressed as

$$\phi(t) = 2\pi \left(-\frac{BW}{2t_{DM}} t^2 + \frac{BW}{2} t \right), \quad t \in [0, t_{DM}] \quad (4)$$

Finally, the time representation of the pulse shape within its period number k is expressed by:

$$v_k(t) = A_k(t) e^{j(\phi(t) + \theta_k)} \text{ for } t \in [0, t_{DM}]$$

$$v_k(t) = 0 \text{ otherwise} \quad (5)$$

where $\phi(t)$ is defined by Equ. (4), T_p is the pulsar period, θ_k is a random phase and $A_k(t)$ is a random amplitude which represents the scintillation effects. Indeed, inhomogeneities in the ISM cause multi-path propagation which produces interference of the signals and diffraction pattern. As a result, we observe changes in intensity with both frequency and time. The quadratic term in Equ. (4) is directly related to the dispersion function defined by Equ. (1) if we consider also linear frequency shift over the bandwidth (i.e. $f_0 \gg \Delta f$). So, we will use the same de-dispersion algorithm as for real pulsar signal.

Figure 4 shows an example of a generated pulsar and the result after de-dispersion. Results are the one expected. With this model, we can generate all kind of scenario to test our RFI mitigation algorithms. Next subsection will present a first algorithm based on the cyclostationary properties of RFI.

4. RFI CYCLOSTATIONNARY DETECTOR AND THE INTEREST OF BLANKING

The proposed detector is based on RFI specific properties named cyclostationarity. Indeed, most of telecommunications signals present a hidden periodicity which is usually scrambled by the intrinsic signal randomness [5,6]. For example, this hidden periodic characteristic can be generated by the carrier frequency or the baud rate of the incoming RFI.

3.4 Cyclostationary detector principle

In [7], we have defined and study the statistics of a cyclostationary detector based on the following criteria:

$$C_{s(*)}^\alpha = \left| \frac{1}{N} \sum_{n=0}^N s(n) s(n)^* e^{-2\pi\alpha n} \right| \quad (6)$$

Where α is the frequency related to the hidden periodicity (α is also called the cyclic frequency), $s(n)$ is a signal to be controlled and $*$ is the conjugate operator which may be used or not. C_s^α and C_{s*}^α are called respectively the conjugate cyclostationary detector and the cyclostationary detector.

Asymptotically, if $s(n)$ is stationary, $C_{s(*)}^\alpha$ will be null. Otherwise, if $s(n)$ is cyclostationary with cyclic frequency α , then $C_{s(*)}^\alpha$ will be non zero. Notice that C_{s*}^0 is simply a power detector.

We have also defined normalized detection criterion:

$$D_{s(*)}^\alpha = \frac{C_{s(*)}^\alpha}{C_{s*}^0} \quad (7)$$

If the cyclic frequency to be detected is not known, we can implement blind cyclic detectors by testing all cyclic frequencies. This can be easily done through FFTs. If the cyclic frequency does not map strictly the FFT bins, the detectors will lose a maximum of 3db in term of sensitivity [7].

In practice, the detectors are applied on a sliding temporal window of N samples. If detection occurs, the whole N samples are flagged as polluted. Depending on the RFI strategy, these N samples can be blanked (i.e. replace by 0) or not.

3.5 Simulation and results

For the simulation, we have considered a BPSK (Binary Phase Shift keying) RFI:

$$rfi(t) = \sum_k a_k h(t - kT_d) e^{i(2\pi f_c t + \varphi)} \quad (8)$$

where f_c is the carrier frequency, $1/T_d$ is the baud rate, a_k is a random binary signal, $h(t)$ is the emission filter and φ is a random phase.

Theoretically, in the case, the cyclic frequencies are $\alpha = n/T_d$, $n \in \mathbb{Z}$ for the $C_{s^*}^\alpha$ detector and $\alpha = 2f_c + n/T_d$, $n \in \mathbb{Z}$ for the C_s^α detector. Figure 5 shows the detector values as a function of the cyclic frequency. In this example, we retrieve the above theoretical cyclic frequencies. It shows also that the conjugate cyclostationary detector with $\alpha = 2f_c$ is the most contrasted criteria.

The generated signal will have the following structure:

$$s(t) = \sum_k v_k(t - kT_p) + rfi(t) + n(t) \quad (9)$$

Where $v_k(t)$ is defined by Eq. (7), T_p is the pulsar period, $rfi(t)$ is defined by Eq. (8), and $n(t)$ is a complex Gaussian noise.

Let us also define the interference to noise ratio

$$INR_{dB} = 10 \log_{10} \left(\frac{P_{rfi}}{P_{noise}} \right) \quad (10)$$

Where, P_{rfi} and P_{noise} are the total power of respectively the RFI and system noise.

In addition, we define the pulse signal to noise ratio as

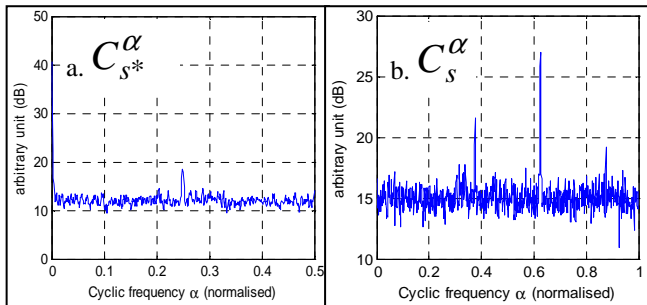


Figure 5 – Cyclic frequency measurements on the simulated BPSK signal with the following parameters $f_c = 0.3125 F_s$ and $T_d = 4/F_s$ (F_s is the sampling frequency). The INR_{dB} has been set to 5 dB. **a.** the cyclostationary detector $C_{s^*}^\alpha$. **b.** the conjugate cyclostationary detector C_s^α . The conjugate cyclostationary detector with $\alpha = 2f_c$ is the most contrasted criteria.

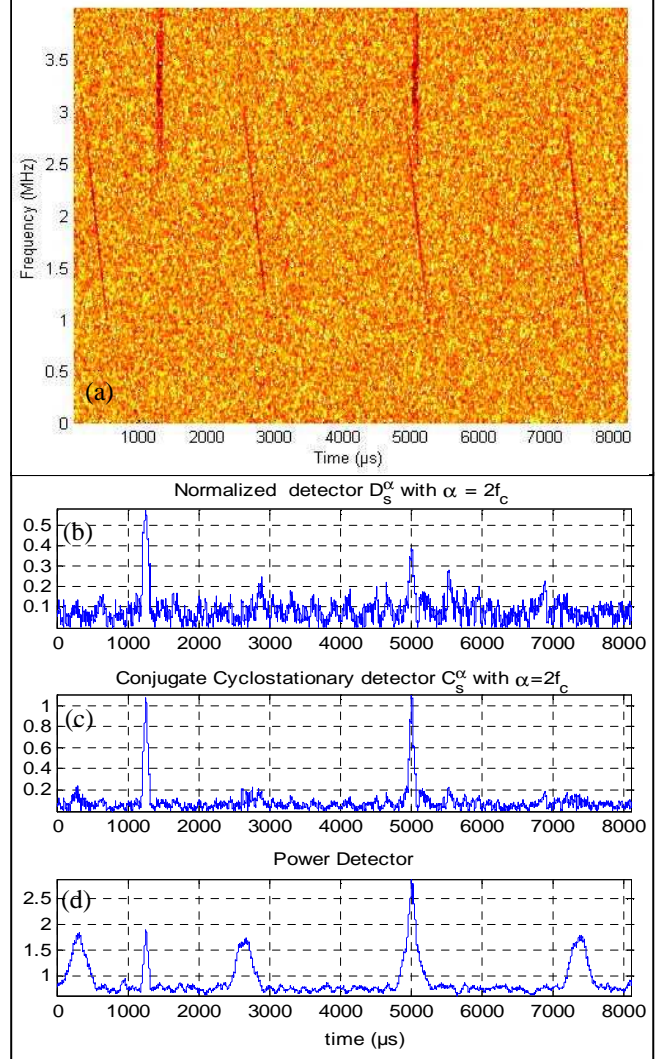


Figure 6 – Detection of impulsive or burst broadband RFI signals by using the conjugate cyclostationary detector **a.** Time-frequency representation of the pulsar signal generated by the model given by Equ. 7 ($BW=4\text{MHz}$, $t_{DM}=938.0793 \mu s$ and pulsar period is 2.35 ms, $RSB= -1 \text{ dB}$) and 2 short BPSK RFI signals generated with Equ. 10 (duration= $80\mu s$, $f_c=25 \text{ kHz}$, $T_d=0.5\mu s$, $h(t)$ is root raised cosine filter with roll-off equal to 1 $INR=3\text{dB}$). **b.** the normalized conjugate cyclostationary detector over time ($N=256$ samples), **c.** The conjugate cyclostationary detector ($N=256$ samples). **D.** the Power detector ($N=256$ samples).

$$SNR_{dB} = 10 \log_{10} \left(\frac{P_{Pulses}}{P_{noise}} \right) \quad (11)$$

Where, P_{pulses} is the total power of the Pulsar's pulses.

The results, in figure 6, show that the cyclostationary detector detects only the bursts. While the power detector cannot make the difference between bursts and Pulsar's pulses.

The next step is to decide to eliminate the RFI or not, taking into account the impact of this decision on the quality of the waveform of the pulsar, or other characteristic which may be distorted. See Figure 7.

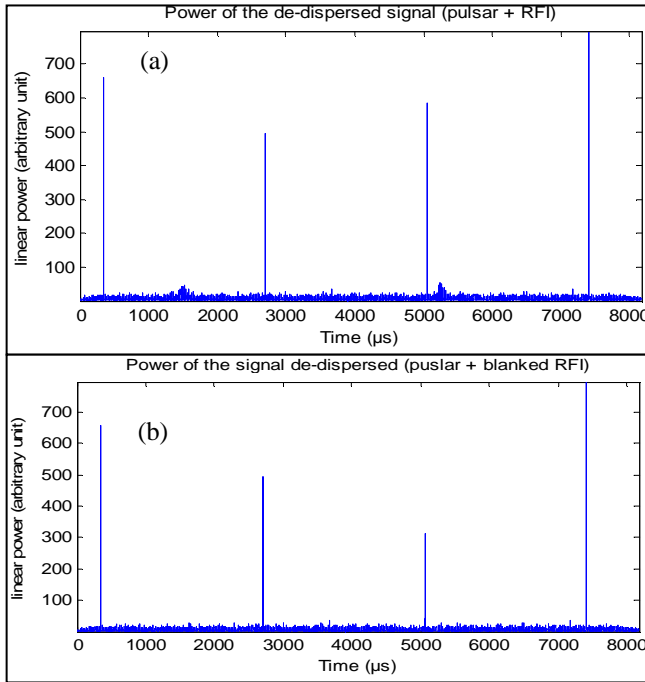


Figure 7 – Blanking of impulsive or burst broadband RFI signals. **b.** Temporal power fluctuations over the whole bandwidth without detection and blanking. **c.** Temporal power fluctuations over the whole bandwidth with detection and blanking. The RFI signals have been removed. Since a part of the 3rd pulse energy is also blanked, the power level of this 3rd pulse is modified.

Depending on where these detectors will be implemented in the system, different type of RFI can be detected. Three possibilities have been identified on the system architecture defined on Figure 2:

(A) at the input of the polyphase filter bank, just after the digitalisation. This configuration is appropriate for impulsive or burst broadband RFI. Figure 6 and Figure 7 shows a simulated example of detection and blanking in such configuration. The power detector will detect pulsar pulses and RFI pulses. The conjugate cyclic detector will only detect the RFI.

(B) Just after the polyphase filter bank.

(C) Just after the FFT, in the coherent de-dispersion process.

This configuration is appropriate for narrow band and continuous RFI.

5. CONCLUSIONS

In the context of pulsar observation, a processing architecture based on 4 Graphic Processing Unit (GPU board) has been presented. This system performs the real time coherent de-dispersion process. It is now fully operational at Nançay Observatory and it outperforms the performances of the previous system based on a cluster of 77 bi-processor Athlon 1.2 GHz. The next step will be to integrate real time RFI detectors. By simulation, we have shown that cyclostationary detectors can be a very interesting alternative to power detectors.

REFERENCES

- [1] D. Lorimer and M. Kramer, *Handbook of Pulsar Astronomy*. Cambridge, United Kingdom: Cambridge University Press, 2005.
- [2] Paul B. Demorest, "Measuring the Gravitational Wave Background using Precision Pulsar Timing," *thesis*, Chapter 3, University of California, Berkeley. Spring. 2007.
- [3] <http://www.cv.nrao.edu/~pdemores/gpu/>
- [4] <http://www.obs-nancay.fr/>
- [5] E. Serpedin, F. Panduru, I. Sari, and G.B. Giannakis, "Bibliography on cyclostationarity," *Signal Processing*, vol. 85, Dec. 2005, pp. 2233-2303.
- [6] W.A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Processing*, vol. 86, Apr. 2006, pp. 639-697
- [7] R. Weber, P. Zarka, V. B. Ryabov, R. Feliachi, J.M. Grießmeier, L. Denis, R. V. Kozhyn, V. V. Vinogradov, P. Ravier, "Data pre-processing for decametre wavelength exoplanet detection: An example of cyclostationary RFI detector," in *Proc. EUSIPCO 2007*, Poznan, Poland, September 3-7. 2007.

Résumé :

The work presented in this thesis is in the context of the intermittent impulsive event detection at Nançay Observatory. The pulsars are highly magnetized neutron stars in rapid rotation, which emit a radio beam scanning the space like a lighthouse. They are detectable with a specific instrumentation. In recent years, new classes of such pulsars were discovered. These pulsars with extreme features, especially with individual pulses more intense and irregular compared to the average, must be detected in real time in a disrupted radio environment because of telecommunication signals. This study presents some radio frequency interference (RFI) mitigation algorithms adapted to this context. Several methods are presented and compared. Among them, two were selected and compared using Monte Carlo simulations with a set of parameters to simulate the pulsar and a BPSK signal with power and different durations. In the case of researching new pulsars, an alternative method is proposed (SIPSFAR), combining research capacity in real time and robustness against RFI. It is based on 2D Fourier transform and the Radon transform. A theoretical comparative study has confronted and compared the sensitivity of this new method and the commonly method used by radio astronomers. SIPSFAR was implemented on a GPU GTX285 and tested on a large survey of the sky made at Nançay radio telescope. The results have led to a further statistical comparison from the actual data.

Mots clés : RFI, detection, cyclostationnarity, Radon, GPU, pulsar

Impulsive event detection in a disturbed radio environment. Application to the observation of intermittent pulsars with real-time signal processing system

Abstract :

Les travaux présentés dans ce mémoire s'inscrivent dans le cadre de la détection d'événements impulsionnels intermittents en provenance de pulsars. Ces objets astrophysiques sont des étoiles à neutrons hautement magnétisées en rotation rapide, qui émettent un faisceau radio balayant l'espace comme la lentille d'un phare. Ils sont détectables grâce à une instrumentation spécifique. Depuis quelques années, on a découvert de nouvelles catégories de ces pulsars, aux caractéristiques extrêmes, avec en particulier des impulsions individuelles plus intenses et irrégulières comparé à la moyenne. Il faut pouvoir les détecter en temps réel dans un environnement radio perturbé à cause des signaux de télécommunications. Cette étude propose des algorithmes de traitement d'interférences radio fréquence (RFI) adaptés à ce contexte. Plusieurs méthodes de traitement de RFI sont présentées et comparées. Parmi elles, deux ont été retenues et comparées au moyen de simulations Monte Carlo, avec un jeu de paramètres simulant le pulsar et un signal BPSK avec des puissances et des durées différentes. Pour la recherche de nouveaux pulsars, une méthode alternative est proposée (SIPSFAR), combinant capacité de recherche en temps réel et robustesse contre les RFI. Elle est basée sur la transformée de Fourier 2D et la transformée de Radon. Une étude comparative théorique a permis de confronter et comparer la sensibilité de cette nouvelle méthode avec celle communément utilisée par les radioastronomes. La méthode a été implantée sur un GPU GTX285 et testée sur un grand relevé du ciel effectué au radiotélescope de Nançay. Les résultats obtenus ont donné lieu à une comparaison statistique complémentaire à partir de données réelles.

Keywords : RFI mitigation, detection, cyclostationnarity, Radon, GPU, pulsar